

# BUNDESREPUBLIK DEUTSCHLAND

09.02.2004

**PRIORITY DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)



## Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**Aktenzeichen:** 103 41 272.7

**Anmeldetag:** 08. September 2003

**Anmelder/Inhaber:** BASF Aktiengesellschaft, 67056 Ludwigshafen/DE

**Bezeichnung:** Verfahren zur gentechnischen Veränderung von Organismen der Gattung *Blakeslea*, entsprechende Organismen und deren Verwendung

**IPC:** C 12 N 15/80

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 29. Januar 2004  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

Hintermeier

**Verfahren zur gentechnischen Veränderung von Organismen der Gattung**  
**Blakeslea, entsprechende Organismen und deren Verwendung**

Die Erfindung betrifft ein Verfahren zur gentechnischen Veränderung von Organismen der Gattung Blakeslea, entsprechende Organismen und deren Verwendung.

Pilze der Gattung Blakeslea sind als Produktionsorganismen bekannt. So wird z. B. Blakeslea trispora als Produktionsorganismus für  $\beta$ -Carotin (Ciegler, 1965, Adv Appl Microbiol. 7:1) und Lycopin verwendet (EP 1201762, EP 1184464, WO 03/038064). Daneben kommt Blakeslea zur Produktion anderer lipophiler Substanzen in Frage wie z.B. andere Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder zur Produktion hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitamine, Provitamine und Cofaktoren.

Die hohen Produktivitäten für  $\beta$ -Carotin und Lycopin machen Blakeslea, insbesondere Blakeslea trispora attraktiv für die wirtschaftliche fermentative Herstellung von Carotinoiden und deren Vorstufen.

Allerdings ist es auch von Interesse die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle zu ermöglichen, die von Blakeslea bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Carotinoide werden Futtermitteln, Nahrungsmitteln, Nahrungsergänzungsmitteln, Kosmetika und Arzneimitteln zugesetzt. Die Carotinoide dienen vor allem als Pigmente zur Färbung. Daneben werden die antioxidative Wirkung der Carotinoide und andere Eigenschaften dieser Substanzen genutzt. Man unterteilt die Carotinoide in die reinen Kohlenwasserstoffe, die Carotine und die sauerstoffhaltigen Kohlenwasserstoffe, die Xanthophylle. Xanthophylle wie Canthaxanthin und Astaxanthin werden beispielsweise zur Pigmentierung von Hühner-

eiem und Fischen eingesetzt (Britton et al. 1998, Carotinoids, Vol 3, Biosynthesis and Metabolism). Die Carotine  $\beta$ -Carotin und Lycopin werden vor allem in der Humanernährung eingesetzt.  $\beta$ -Carotin wird beispielsweise als Getränkefarbstoff verwendet. Lycopin hat eine krankheitsvorbeugende Wirkung (Argwal und Rao, 2000, CMAJ 163:739-744; Rao und Argwal 1999, Nutrition Research 19:305-323). Die farblose Carotinoidvorstufe Phytoen kommt vor allem für Anwendungen als Antioxidans in Frage.

Der überwiegende Teil der Carotinoide und deren Vorstufen die als Zusatzstoffe für die oben genannten Anwendungen eingesetzt werden, wird durch chemische Synthese hergestellt. Die chemische Synthese ist mehrstufig, technisch sehr aufwendig und verursacht hohe Herstellkosten. Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten.

Hierzu ist eine gentechnische Veränderung von *Blakeslea* erforderlich. Insbesondere, wenn Xanthophylle produziert werden sollen, da diese Verbindungen natürlicherweise vom Wildtyp der *Blakeslea* nicht synthetisiert werden.

Allerdings sind bisher keine Methoden zur gentechnischen Veränderung von *Blakeslea*, insbesondere *Blakeslea trispora* bekannt.

Als Methode zur Herstellung von gentechnischen veränderten Pilzen wurde in einigen Fällen die *Agrobacterium*-vermittelte Transformation erfolgreich eingesetzt. So sind z. B. folgende Organismen durch *Agrobakterien* transformiert worden: *Saccharomyces cerevisiae* (Bundock et al., 1995, EMBO Journal, 14:3206-3214), *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Colletotrichum gloeosporioides*, *Fusarium solani pisi*, *Neurospora crassa*, *Trichoderma reesei*, *Pleurotus ostreatus*, *Fusarium graminearum* (van der Toorren

et al., 1997, EP 870835), *Agraricus bisporus*, *Fusarium venenatum* (de Groot et al., 1998, *Nature Biotechnol.* 16:839–842), *Mycosphaerella graminicola* (Zwiers et al. 2001, *Curr. Genet.* 39:388–393), *Glarea lozoyensis* (Zhang et al., 2003, *Mol. Gen. Genomics* 268:645–655), *Mucor miehei* (Monfort et al. 2003, *FEMS Microbiology Lett.* 244:101 – 106).

Von Interesse ist besonders eine homologe Rekombination, bei der zwischen der einzuführenden DNA und der Zell-DNA möglichst viele Sequenzhomologien bestehen, so dass eine ortsspezifische Einführung bzw. Ausschaltung von genetischer Information im Genom des Empfängerorganismus möglich ist. Andernfalls wird die Spender-DNA durch illegitime bzw. nicht-homologe Rekombination ins Genom des Empfängerorganismus integriert, was nicht ortsspezifisch erfolgt.

Eine durch *Agrobacterium* vermittelte Transformation und anschließende homologe Rekombination der transferierten DNA wurde bisher bei folgenden Organismen nachgewiesen: *Aspergillus awamori* (Gouka et al. 1999, *Nature Biotech* 17:598-601), *Glarea lozoyensis* (Zhang et al., 2003, *Mol. Gen. Genomics* 268:645-655), *Mycosphaerella graminicola* (Zwiers et al. 2001, *Curr. Genet.* 39:388-393).

Als weitere Methode zur Transformation von Pilzen ist die Elektroporation bekannt. Die integrative Transformation von Hefe durch Elektroporation wurde von Hill, *Nucl. Acids. Res.* 17:8011 gezeigt. Für filamentöse Pilze wurde die Transformation durch Chakaborty und Kapoor beschrieben (1990, *Nucl. Acids. Res.* 18:6737).

Eine „biolistische“ Methode, d.h. die Übertragung von DNA durch Beschuss von Zellen mit DNA-beladenen Partikeln wurde beispielsweise für *Trichoderma harzianum* und *Gliocladium virens* beschrieben (Lorito et al. 1993, *Curr. Genet.* 24:349–356).



Diese Methoden konnten bisher jedoch nicht erfolgreich zur gezielten genetischen Veränderung von *Blakeslea* und insbesondere *Blakeslea trispora* eingesetzt werden.

- 5 Eine besondere Schwierigkeit bei der Herstellung von gezielt genetisch veränderten *Blakeslea* und *Blakeslea trispora*, ist die Tatsache, dass deren Zellen in allen Stadien des sexuellen und des vegetativen Zellzyklus mehrkernig sind. In Sporen von *Blakeslea trispora* Stamm NRRL2456 und NRRL2457 wurden z. B. im Durchschnitt 4,5 Kerne pro Spore nachgewiesen (Metha und Cerdá-Olmedo, 10 1995, Appl. Microbiol. Biotechnol. 42:836-838). Dies hat zur Folge, dass die gentechnische Veränderung in aller Regel nur in einem oder wenigen Kernen vorliegt, die Zellen also heterokaryotisch sind.

- 15 Sollen die genetisch veränderten *Blakeslea*-Arten, insbesondere *Blakeslea trispora* zur Produktion eingesetzt werden, so ist es insbesondere bei einer Gendeletion wichtig, dass in den Produktionsstämmen die gentechnische Veränderung in allen Kernen vorliegt, so dass eine stabile und hohe Syntheseleistung ohne Nebenprodukte möglich wird. Die Stämme müssen folglich in Bezug auf die gentechnische Veränderung homokaryotisch sein.

- 20 Lediglich für *Phycomyces blakesleeanus* ist ein Verfahren beschrieben worden, um homokaryotische Zellen zu erzeugen (Roncero et al., 1984, Mutat. Res. 125:195). Durch Zugabe des mutagenen Agens MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) werden nach dem dort beschriebenen Verfahren Kerne in den 25 Zellen eliminiert, so dass statistisch eine gewisse Anzahl von Zellen mit nur noch einem funktionellem Kern vorliegt. Die Zellen werden dann einer Selektion unterzogen, in der nur einkernige Zellen mit einem rezessiven Selektionsmarker zu einem Mycel auswachsen können. Die Nachkommen dieser selektierten Zellen sind mehrkernig und homokaryotisch. Ein rezessiver Selektionsmarker für 30 *Phycomyces blakesleeanus* ist z. B. dar.  $Dar^+$ -Stämme nehmen das toxische Riboflavin-Analog 5-Carbon-5-deazariboflavin auf;  $Dar^-$ -Stämme dagegen nicht (Delbrück et al. 1979, Genetics 92:27). Die Selektion von rezessiven Mutanten erfolgt durch Zugabe von 5-Carbon-5-deazariboflavin (DARF).

Allerdings ist dieses Verfahren nicht für *Blakeslea*, insbesondere *Blakeslea trispora* bekannt und insbesondere nicht mit im Zusammenhang mit einer Transformation beschrieben worden.

- 5 Aufgabe der vorliegenden Erfindung ist es demgegenüber, ein Verfahren bereitzustellen, mit dem eine gentechnische Veränderung von *Blakeslea*-Stämmen, insbesondere *Blakeslea trispora* möglich ist. Darüber hinaus ist es Aufgabe der Erfindung ein Verfahren bereitzustellen, das die Herstellung homokaryotischer genetisch veränderter Stämme erlaubt. Ferner ist es eine Aufgabe
- 10 der Erfindung entsprechend gentechnisch veränderte Zellen bereitzustellen.

Diese Aufgabe wird durch ein Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* gelöst, umfassend

- (i) Transformation mindestens einer der Zellen,
- 15 (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen durch Eliminierung von Kernen in diesen Zellen, so dass mindestens eine Zelle mit nur einem Kern verbleibt, der die gewünschte genetische Veränderung aus (i) aufweist und zur Ausprägung führt, und
- (iii) Selektion der gentechnisch veränderten Zelle oder Zellen.

20

Mit der erfindungsgemäßen Methode ist es möglich, mehrkernige Zellen der Pilze *Blakeslea* gezielt und stabil genetisch zu verändern, um so Mycel aus Zellen mit einheitlichen Kernen zu gewinnen. Vorzugsweise handelt es sich um Zellen von Pilzen der Art *Blakeslea trispora*

25

Unter Transformation wird die Übertragung einer genetischen Information in den Organismus, insbesondere Pilz verstanden. Darunter sollen alle dem Fachmann bekannten Möglichkeiten zur Einschleusung der Information, insbesondere DNA fallen, z. B. Beschuss mit DNA-beladenen Partikeln, Transformation mittels Protoplasten, Mikroinjektion von DNA, Elektroporation, Konjugation

30 oder Transformation kompetenter Zellen, Chemikalien oder Agrobakterien ver-

mittelte Transformation. Als genetische Information werden ein Genabschnitt, ein Gen oder mehrere Gene verstanden. Die genetische Information kann z. B. mit Hilfe eines Vectors oder als freie Nukleinsäure (z. B. DNA, RNA) und auf sonstige Weise in die Zellen eingebracht und entweder durch Rekombination  
5 ins Wirtsgenom eingebaut oder in freier Form in der Zelle vorliegen. Besonders bevorzugt ist hierbei die homologe Rekombination.

Bevorzugte Transformationsmethode ist die *Agrobacterium tumefaciens*-vermittelte Transformation. Hierzu wird zunächst die zu transferierende Spen-  
10 der-DNA in einen Vektor eingefügt, der (i) flankierend zu der zu transferierenden DNA die T-DNA-Enden trägt, der (ii) einen Selektionsmarker enthält und der (iii) ggf. Promotoren und Terminatoren für die Genexpression der Spender-DNA aufweist. Dieser Vektor wird in einen *Agrobacterium tumefaciens*-Stamm übertragen, der ein Ti-Plasmid mit den vir-Genen enthält. vir-Gene sind für den  
15 DNA-Transfer in *Blakeslea* verantwortlich. Mit diesem Zwei-Vektor-System wird die DNA von *Agrobacterium* in *Blakeslea* übertragen. Hierzu werden die *Agrobakterien* zunächst in Gegenwart von Acetosyringone inkubiert. Acetosyringone induziert die vir-Gene. Anschließend werden Sporen von *Blakeslea trispora* zusammen mit den induzierten Zellen von *Agrobacterium tumefaciens* auf Aceto-  
20 syringone-haltigem Medium inkubiert und dann auf Medium übertragen, das eine Selektion der Transformanten, d.h. der gentechnisch veränderten Stämme von *Blakeslea* ermöglicht.

25 Der Begriff Vector wird in der vorliegenden Anmeldung als eine Bezeichnung für ein DNA-Molekül verwendet, das zum Einschleusen und ggf. zur Vermehrung von Fremd-DNA in eine Zelle dient (siehe auch "Vector" in Römpp Lexikon Chemie – CDROM Version 2.0, Stuttgart/New York: Georg Thieme Verlag 1999). In der vorliegenden Anmeldung sollen unter dem begriff "Vector" auch  
30 Plasmide, Cosmide usw. verstanden werden, die dem gleichen Zweck dienen.

Unter Expression wird in der vorliegenden Anmeldung die Übertragung einer genetischen Information ausgehend von DNA oder RNA in ein Gen-Produkt (hier vorzugsweise Carotinoide) verstanden und soll auch den Begriff der Über-

expression beinhalten, womit eine verstärkte Expression gemeint ist, so dass das ein bereits in der nicht transformierten Zelle (Wildtyp) hergestelltes Produkt verstärkt produziert wird oder einen großen Teil des gesamten Gehaltes der Zelle ausmacht.

5

Unter gentechnischer Veränderung soll die Einschleusung genetischer Information in einen Empfängerorganismus, so dass diese stabil exprimiert und bei der Zellteilung weitergegeben wird, verstanden werden. In diesem Zusammenhang ist die Homokaryotisierung, die Herstellung von Zellen, die nur einheitliche Kerne enthalten, d. h. Kerne mit gleichem genetischem Informationsgehalt.

10

Diese Homokaryotisierung ist nur notwendig, wenn die durch Transformation eingeführte genetische Information rezessiv vorliegt, d. h. nicht zur Ausprägung gelangt. Führt die Transformation aber zu einem dominanten Vorliegen der genetischen Information, d. h. wird sie ausgeprägt, so ist eine Homokaryotisierung nicht unbedingt nötig. Vorzugsweise wird zur Homokaryotisierung ein mutagenes Agens eingesetzt, wobei es sich insbesondere um N-Methyl-N'-nitro-nitrosoguanidin (MNNG) handelt.

15

Unter Selektion wird die Auswahl von Zellen verstanden, deren Kerne dieselbe genetische Information beinhalten, d. h. Zellen die die gleichen Eigenschaften aufweisen, wie Resistenzen oder die Herstellung bzw. vermehrte Herstellung eines Produktes. In der Selektion werden bevorzugt 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) eingesetzt.

20

25

Der in der Transformation (i) eingesetzte Vector kann derart gestaltet sein, dass die im Vector enthaltene genetische Information in das Genom mindestens einer Zelle integriert wird. Dabei kann genetische Information in der Zelle ausgeschaltet werden.

30

Der in der Transformation (i) eingesetzte Vector kann aber auch derart ausgestaltet sein, dass die im Vector enthaltene genetische Information in der Zelle exprimiert wird, d. h. genetische Information eingefügt wird, die im korrespon-

dierenden Wildtyp nicht vorhanden ist oder die durch die Transformation verstärkt bzw. überexprimiert wird.

5 Der Vector kann beliebige genetische Informationen zur genetischen Veränderungen von Organismen der Gattung *Blakeslea* enthalten.

10 Unter „genetischer Information“ werden vorzugsweise Nukleinsäuren verstanden, deren Einbringung in den Organismus der Gattung *Blakeslea* zu einer genetischen Veränderung in Organismen der Gattung *Blakeslea*, also beispielsweise zu einer Verursachung, Erhöhung oder Reduzierung von Enzymaktivitäten im Vergleich zum Ausgangsorganismus führen.

15 Der Vector kann beispielsweise genetische Information zur Herstellung lipophiler Substanzen enthalten wie z.B. Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder genetische Information zur Herstellung hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitaminen, Provitamine und Cofaktoren.

20 Bevorzugterweise enthält der eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder Xanthophyllen oder deren Vorstufen.

25 Besonders bevorzugt sind genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon,  $\beta$ -Cryptoxanthin, Anthonixanthin, Adonirubin, Canthaxanthin, 3- und 3'-Hydroxyechinenon, Lycopin, Lutein,  $\beta$ -Carotin oder Phytoen. Ganz besonders bevorzugt sind genetische Informationen zur Herstellung von Phytoen, Astaxanthin oder Zeaxanthin.

30 In einer bevorzugten Ausführungsform werden die Organismen der Gattung *Blakeslea* beispielsweise dadurch in die Lage versetzt Xanthophylle, wie beispielsweise Canthaxanthin, Zeaxanthin oder Astaxanthin herzustellen, indem in den genetisch veränderten Organismen der Gattung *Blakeslea* im Vergleich zum Wildtyp eine Hydroxylase-Aktivität und/oder Ketolase-Aktivität verursacht wird.

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Canthaxanthin umzuwandeln.

Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge  $\beta$ -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende nicht genetisch veränderte Ausgangsorganismus der Gattung *Blakesleaa* verstanden.

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) der Gattung *Blakesleaa* oder ein erfindungsgemäßer, genetisch veränderter Organismus der Gattung *Blakesleaa* oder beides verstanden werden.

Vorzugsweise wird unter "Wildtyp" für die Verursachung der Ketolase-Aktivität und für die Verursachung der Hydroxylase-Aktivität jeweils ein Referenzorganismus verstanden.

Dieser Referenzorganismus der Gattung *Blakesleaa* ist *Blakeslea trispora* ATCC 14271.

Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakesleaa* und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Bestimmung der Ketolase-Aktivität in Organismen der Gattung *Blakeslea* erfolgt in Anlehnung an die Methode von Frazer et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels HPLC ermittelt.

Der erfindungsgemäße genetisch veränderte Organismus der Gattung *Blakeslea* weist in dieser, bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren.

In einer weiter bevorzugten Ausführungsform erfolgt die Verursachung der Ketolase-Aktivität in den Organismen der Gattung *Blakeslea* durch Genexpression einer Nukleinsäure kodierend eine Ketolase.

In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren in den Ausgangsorganismus der Gattung *Blakeslea*.

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäuren die eine Ketolase codiert verwendet werden.

Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus der Gattung *Blakeslea* nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

Beispiele für Nukleinsäuren, kodierend eine Ketolase und die entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren verwendet werden können sind beispielsweise Sequenzen aus:

5. *Haematoccus pluvialis*, insbesondere aus *Haematoccus pluvialis* Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 11, Protein SEQ ID NO: 12),

- 10 *Haematoccus pluvialis*, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 13, Protein SEQ ID NO: 14),

*Agrobacterium aurantiacum* (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

- 15 *Alicycobaculum spec.* (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),

*Paracoccus marcusii* (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20).

20 *Synechocystis* sp. Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22).

- 25 *Bradyrhizobium* sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).

*Nostoc* sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26),

- 30 *Nostoc punctiforme* <sup>C</sup> ATTC 29133, Nukleinsäure: Acc.-No. NZ\_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 27); Protein: Acc.-No. ZP\_00111258 (SEQ ID NO: 28) (als putatives Protein annotiert),



*Nostoc punctiforme* ATTC 29133, Nukleinsäure: Acc.-No. NZ\_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 29), Protein: (SEQ ID NO: 30) (nicht annotiert).

- 5 Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit den vorstehend
- 10 beschriebenen Sequenzen und insbesondere mit den Sequenzen SEQ ID NO: 12 und/oder 26 leicht auffinden.

- Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen,
- 15 insbesondere ausgehend von den Sequenzen SEQ ID NO: 12 und/oder 26 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

- Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.
- 20

- Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6
- 25 beschrieben.

- Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer
- 30 Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0.2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

Darüberhinaus kann die Temperatur während des Waschschrilles von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

- 5 Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50 % Formamid wird die Hybridisierung bevorzugt bei 42°C
- 10 ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind infolge gegeben:

- 15 (1) Hybridisierungsbedingungen mit zum Beispiel
- (i) 4X SSC bei 65°C, oder
  - (ii) 6X SSC bei 45°C, oder
  - (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
  - (iv) 6X SSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA bei 68°C, oder
  - 20 (v) 6XSSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder
  - (vi) 50 % Formamid, 4X SSC bei 42°C, oder
  - (vii) 50 % (vol/vol) Formamid, 0.1 % Rinderserumalbumin, 0.1 % Ficoll, 0.1 %
  - 25 Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6.5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder
  - (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
  - (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Bedingungen).
  - 30
- (2) Waschschrille für jeweils 10 Minuten mit zum Beispiel
- (i) 0.015 M NaCl/0.0015 M Natriumcitrat/0.1 % SDS bei 50°C, oder
  - (ii) 0.1X SSC bei 65°C, oder
  - (iii) 0.1X SSC, 0.5 % SDS bei 68°C, oder

- (iv) 0.1X SSC, 0.5 % SDS, 50 % Formamid bei 42°C, oder
- (v) 0.2X SSC, 0.1 % SDS bei 42°C, oder
- (vi) 2X SSC bei 65°C (moderate Bedingungen).

5 In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakeslea* bringt man Nukleinsäuren ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 12 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 und die enzymatische Eigenschaft einer Ketolase aufweist.

15

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 12 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

20

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 26 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 26 und die enzymatische Eigenschaft einer Ketolase aufweist.

25

30

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen

Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 26 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

5

Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sog. konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähnliche Eigenschaft hat wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

10

Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptids und die Verknüpfungen zwischen den einzelnen Proteindomänen.

15

Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren ersetzt wird.

20

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden, insbesondere die Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma DNASTAR, inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

25

Multiple alignment parameter:

Gap penalty 10

30

Gap length penalty 10

Pairwise alignment parameter:

K-tuple 1

Gap penalty 3

Window 5

Diagonals saved 5

Unter einem Protein, das eine Identität von mindestens 20 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 oder 26 aufweist, wird dementsprechend ein Protein verstanden, das bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12 oder 26, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 20 % aufweist.

10 Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Blakesleaspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene von Organismen der Gattung Blakeslea leicht ermitteln.

20 In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 11 in die Organismus der Gattung ein.

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 25 in die Organismus der Gattung ein.

25 Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der  
30 Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al.

(1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

- 5 Unter Hydroxylase-Aktivität die Enzymaktivität einer Hydroxylase verstanden.

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten,  $\beta$ -Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

10

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist,  $\beta$ -Carotin in Zeaxanthin oder Cantaxanthin in Astaxanthin umzuwandeln.

- 15 Dementsprechend wird unter Hydroxylase-Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge  $\beta$ -Carotin oder Cantaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

- 20 Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein Hydroxylase die umgesetzte Menge  $\beta$ -Carotin oder Cantaxanthin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

- 25 Vorzugsweise beträgt diese Erhöhung der Hydroxylase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der Hydroxylase-Aktivität des Wildtyps.

- 30 Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenz-Organismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) *in vitro* bestimmt. Es wird zu einer bestimmten Menge an Organismenextrakt Ferredoxin, Ferredoxin-NADP Oxidoreductase, Katalase, NADPH sowie beta-Carotin mit Mono- und Digalaktosylglyzeriden zugegeben.

5

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum* L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

10

Der in-vitro Assay wird in einem Volumen von 0.250 ml Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), 0.025 mg Ferredoxin von Spinat, 0.5 Einheiten Ferredoxin-NADP+ Oxidoreduktase von Spinat, 0.25 mM NADPH, 0.010 mg beta-Carotin (in 0.1 mg Tween 80 emulgiert), 0.05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalase, 200 Mono- und Digalaktosylglyzeriden, (1:1), 0.2 mg Rinderserumalbumin und Organismenextrakt in unterschiedlichem Volumen. Die Reaktionsmischung wird 2 Stunden bei 30°C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.

15

20

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cyclase inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):

25

Der in-vitro Assay wird in einem Volumen von 250 µl Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), unterschiedliche Mengen an Organismenextrakt, 20 nM Lycopin, 250 µg an chromoplastidärem Stroma-protein aus Paprika, 0.2 mM NADP+, 0.2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30°C wird die Reaktion durch Zugabe von Chloro-

30

form/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.

Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und  
5 Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).

Die Erhöhung der Hydroxylase-Aktivität kann durch verschiedene Wege erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch Erhöhung der Genexpression von Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp.  
10

Die Erhöhung der Genexpression der Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp kann ebenfalls durch verschiedene Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehrerer Hydroxylase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure kodierend eine Hydroxylase in denb Organismus der Gattung Blakesleaa.  
15

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure kodierend eine Hydroxylase durch Einbringen von mindestens einer Nukleinsäure kodierend eine Hydroxylase in den Organismus der Gattung Blakesleaa.  
20

Dazu kann prinzipiell jedes Hydroxylase-Gen, also jede Nukleinsäure, die eine Hydroxylase und jede Nukleinsäure, die eine  $\beta$ -Cyclase codiert, verwendet werden.  
25

Bei genomischen Hydroxylase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der WirtsOrganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.  
30



Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, kodierend eine Hydroxylase aus *Haematococcus pluvialis*, Accession AX038729, WO 0061764); (Nukleinsäure: SEQ ID NO: 31, Protein: SEQ ID NO: 32).

5 sowie Hydroxylasen der folgenden Accession Nummern:

|emb|CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1, AF499108\_1, AF315289\_1, AF296158\_1, AAC49443.1, NP\_194300.1, NP\_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1, AAL80006.1, 10 AF162276\_1, AAO53295.1, AAN85601.1, CRTZ\_ERWHE, CRTZ\_PANAN, BAB79605.1, CRTZ\_ALCSP, CRTZ\_AGRAU, CAB56060.1, ZP\_00094836.1, AAC44852.1, BAC77670.1, NP\_745389.1, NP\_344225.1, NP\_849490.1, ZP\_00087019.1, NP\_503072.1, NP\_852012.1, NP\_115929.1, ZP\_00013255.1

15 In den erfindungsgemäßen bevorzugten transgenen Organismen der Gattung *Blakeslea* liegt also in dieser bevorzugten Ausführungsform gegenüber dem Wildtyp mindestens ein weiteres Hydroxylase-Gen vor.

20 In dieser bevorzugten Ausführungsform weist die genetisch veränderte Organismus beispielsweise mindestens eine exogene Nukleinsäure, kodierend eine Hydroxylase oder mindestens zwei endogene Nukleinsäuren, kodierend eine Hydroxylase auf.

25 Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Ausführungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 32 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens 50 %, bevorzugter mindestens 70%, noch bevorzugter mindestens 90 %, am bevorzugtesten mindestens 95 % auf Aminosäureebene mit der Sequenz SEQ. ID. 30 NO: 32, und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz be-

kannt ist, wie vorstehend beschrieben, durch Homologievergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der Seq ID. NO: 32 leicht auffinden.

- 5 Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 31 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

10

In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 32.

15

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

20

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismenspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

25

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 31 in den Organismus ein.

30

Alle vorstehend erwähnten Hydroxylase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und

35 Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in

gationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

- 5 Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise den gpd Promotor und/oder den trpC Terminator. Diese haben sich zur Transformation der *Blakeslea* besonders bewährt.

10 Vorteilhafterweise weist der im Vector eingesetzte gpd Promotor die Sequenz SEQ ID NO: 1 auf. Vorteilhafterweise weist der im Vector eingesetzte trpC Terminator die Sequenz SEQ ID NO: 2 auf.

Insbesondere werden dabei der gpd Promotor und der trpC Terminator aus *Aspergillus nidulans* eingesetzt.

15

Insbesondere enthält der in der Transformation (i) eingesetzte Vector ein Resistenzgen. Bevorzugterweise handelt es sich um ein Hygromycin-Resistenzgen (hph), insbesondere das aus *E. coli*. Dieses Resistenzgen hat sich bei dem Nachweis der Transformation und Selektion der Zellen als besonders geeignet

20

Als Promotor für hph wird also bevorzugt p-gpdA, der Promotor der Glycerinaldehyd-3-phosphatdehydrogenase aus *Aspergillus nidulans* genutzt. Als Terminator für hph wird bevorzugt t-trpC, der Terminator des Gens trpC, codierend für Anthranilatsynthasekomponenten aus *Aspergillus nidulans* genutzt.

25

Der Vector kann beispielsweise die SEQ ID NO: 3 aufweisen.

30 Mit dem erfindungsgemäßen Verfahren sind gentechnisch veränderte Organismen *Blakeslea*, insbesondere der Art *Blakeslea trispora* bzw. aus ihnen gebildetes Mycel erhältlich.

Die genetisch veränderten Organismen können zur Produktion von Carotinoiden, Xanthophyllen oder deren Vorstufen verwendet werden. Auch können

neue, im Wildtyp natürlicherweise nicht vorkommende Carotinoide durch Einbringung der entsprechenden genetischen Information von den gezielt genetisch veränderten Zellen bzw. dem durch sie gebildeten Mycel erzeugt und anschließend isoliert werden.

5

Bevorzugterweise ist die Gewinnung von Carotinoiden oder deren Vorstufen mit den gezielt genetisch veränderten Zellen bzw. das durch sie gebildete Mycel möglich.

10 Die Erfindung wird nachfolgend an Hand von Beispielen näher ausgeführt.

### Beispiele

Molekulargenetische Arbeiten wurden, wenn nicht anders beschrieben, nach den Methoden in Current Protocols in Molecular Biology (Ausubel et al., 1999, John Wiley & Sons) durchgeführt.

15

### Stämme und Wachstumsbedingungen

Der *Blakeslea trispora* Stamm ATCC14272 (-) (ein Wildtyp) wurde erhalten von der American Type Culture Collection. Die Anzucht von *B. trispora* erfolgte in MEP-Medium (Malzextrakt-Pepton-Medium): 30 g/l Malzextrakt (Difco), 3 g/l Pepton (Soytone, Difco), 20 g/l Agar, Einstellung pH 5,5, ad 1000 ml mit H<sub>2</sub>O bei 28 °C.

20

Die Anzucht von *Agrobacterium tumefaciens* LBA4404 erfolgte nach Hoekema et al. (1983, Nature 303:179-180) bei 28 °C für 24 h in Agrobakterien-Minimal Medium (AMM): 10 mM K<sub>2</sub>HPO<sub>4</sub>, 10 mM KH<sub>2</sub>PO<sub>4</sub>, 10 mM Glucose, MM-Salze (2,5 mM NaCl, 2 mM MgSO<sub>4</sub>, 700 µM CaCl<sub>2</sub>, 9 µM FeSO<sub>4</sub>, 4 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>).

25

### Transformation von *Agrobacterium tumefaciens*

Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das A.

30

*tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

### **Transformation von *Blakeslea trispora***

- 5 Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD<sub>600</sub> von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD<sub>600</sub> von ca. 0,6 angezogen.
- 10 Zur Co-Inkubation von *Blakeslea* und *Agrobacterium* wurden 100 µl Agrobaktériensuspension mit 100 µl *Blakeslea* Sporensuspension (10<sup>7</sup> Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt. Zur Selektion auf transformierte *Blakeslea*-zellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l MgSO<sub>4</sub> x 7H<sub>2</sub>O, 150 mg/l KH<sub>2</sub>PO<sub>4</sub>, 25 µg/l ThiaminHCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, pH 5,5, 18 g/l Agar) ausplattiert.
- 20

### **Mutagenese mit MNNG**

- 25 Zur Reduzierung der Anzahl von Kernen pro Spore wurde eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt. Hierfür wurde zunächst eine Sporensuspension mit 1 x 10<sup>7</sup> Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Zeit der Inkubation in
- 30 MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.

### **Selektion homonukleater Zellen**

Die Selektion homonukleater Zellen von *Blakeslea trispora* carB<sup>-</sup> erfolgte analog zum Versuchsprotokoll für *Phycomyces blakesleeanus* (Roncero et al., 1984, Mutation Research, 125:195-204), modifiziert durch Wachstum in Gegenwart von 5-Carbon-5-Deazariboflavin (1 µg/ml) und Hygromycin 100 (µg/ml).

5

### **Herstellung genetisch veränderter *Blakeslea trispora* durch Agrobacterium-vermittelte Transformation**

#### **Herstellung des rekombinanten Plasmids pBinAHyg**

Aus dem Plasmid pANsCos1 (Fig.1, Osiewacz, 1994, Curr. Genet. 26:87-90, SEQ ID NO: 4) wurde die gpdA-hph-trpC-Kassette als BglII/HindIII Fragment isoliert und in das mit BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet (Fig. 2, SEQ ID NO: 3) und enthielt das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors (SEQ ID NO: 1) und des trpC Terminators (SEQ ID NO: 2) aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind.

15

#### **Übertragung von pBinAHyg in *Agrobacterium tumefaciens***

Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

25

#### **Übertragung von pBinAHyg in *Blakeslea trispora***

Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD<sub>660</sub> von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD<sub>660</sub> von ca. 0,6 angezogen.

30

Zur Co-Inkubation von *Blakeslea trispora* (B.t.) und *Agrobacterium tumefaciens* (A.t.) wurden 100 µl Agrobakteriensuspension mit 100 µl *Blakeslea* Sporensus-

pension ( $10^7$  Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt.

5

Zur Selektion auf transformierte Blakeslea-Zellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebil-

10

dete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l  $\text{MgSO}_4 \times 7\text{H}_2\text{O}$ , 150 mg/l  $\text{KH}_2\text{PO}_4$ , 25 µg/l Thiamin-HCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, pH 5,5, 100 mg/l Cefotaxim, 100 mg/l Hygromycin, 18 g/l Agar) ausplattiert. Die Übertragung von Sporen auf frische Selektionsplatten wurde dreimal wiederholt. Auf diese

15

Weise wurde die Transformante *Blakeslea trispora* GVO 3005 isoliert.

#### **Nachweis der genetischen Veränderung von *Blakeslea trispora***

20

200 ml MEP-Medium (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5) wurden mit  $10^5$  bis  $10^7$  Sporen der Transformante *Blakeslea trispora* GVO 3005 beimpft und 7 Tage bei 26 °C mit 200 Upm auf einem Rundschüttler inkubiert. Zum Nachweis der erfolgreichen Transformation wurde DNA aus dem Mycel isoliert (Peglab Fungal DNA Mini Kit) und in einer PCR (Programm: 94 °C 1 min, dann 30 Zyklen mit 1 min. 94 °C, 1 min. 58 °C, 1 min. 72 °C) eingesetzt.

25

Zum Nachweis des Hygromycinresistenzgens (hph) wurden die Primer hph-forward (5'-CGATGTAGGAGGGCGTGGATA, SEQ ID NO: 5) und hph-reverse (5'-GCTTCTGCGGGCGATTTGTGT, SEQ ID NO: 6) verwendet. Das erwartete Fragment von hph wies eine Länge von 800 bp auf.

30

Zur Amplifikation des Kanamycinresistenzgens nptIII und damit als Kontrolle auf Agrobakterien wurden die Primer nptIII-forward (5'-TGAGAATATCACCGGAATTG, SEQ ID NO: 7) und nptIII-reverse (5'-AGCTCGACATACTGTTCTTCC, SEQ ID NO: 8) verwendet. Das erwartete Fragment von nptIII wies eine Länge von 700 bp auf.

Zur Amplifikation eines Fragmentes des Glycerinaldehyd-3-phosphatdehydrogenasegens *gpd1* und damit als Kontrolle auf *Blakeslea trispora* wurden die Primer MAT292 (5'-GTGAATGGAAATCCCATCGCTGTC, SEQ ID NO: 9) und MAT293 (5'-AGTGGGTACTCTAAAGGCCATACC, SEQ ID NO: 10) verwendet. Das erwartete Fragment von *gpd1* wies eine Länge von 500 bp auf.

Das Ergebnis der PCR der *Blakeslea trispora* DNA ist in Fig. 3 anhand eines Standard-Gels gezeigt. Die Spuren des Gels wurden folgendermaßen belegt:

- |    |                                     |  |
|----|-------------------------------------|--|
| 1) | 100 bp Größenmarker (100 bp - 1 kb) |  |
| 2) | B.t. GVO 3005                       | primer <i>nptIII</i> -for / <i>nptIII</i> -rev                           |
| 3) | B.t. GVO 3005                       | primer <i>hph</i> -for / <i>hph</i> -rev                                 |
| 15 | 4)                                  | B.t. GVO 3005 primer MAT292 / MAT293 ( <i>gpd</i> )                      |
|    | 5)                                  | A.t. mit Plasmid pBinAHyg primer <i>nptIII</i> -for / <i>nptIII</i> -rev |
|    | 6)                                  | A.t. mit Plasmid pBinAHyg primer <i>hph</i> -for / <i>hph</i> -rev       |
|    | 7)                                  | B.t. 14272 WT primer <i>nptIII</i> -for / <i>nptIII</i> -rev             |
|    | 8)                                  | B.t. 14272 WT primer <i>hph</i> -for / <i>hph</i> -rev                   |
| 20 | 9)                                  | B.t. 14272 WT primer MAT292 / MAT293 ( <i>gpd</i> )                      |

In der DNA von *Blakeslea trispora* wurde das Hygromycinresistenzgens (*hph*) und als Positivkontrolle Glycerinaldehyd-3-phosphatdehydrogenasegen (*gpd1*) nachgewiesen. *nptIII* konnte demgegenüber nicht nachgewiesen werden.

Somit wurde die genetische Veränderung von *Blakeslea trispora* durch Agrobacterium-vermittelte Transformation nachgewiesen.

#### **Isolierung homokaryotischer GVO von *Blakeslea trispora*:**

Durch erfolgreichen Transfer des Vectors pBinAHyg in *Blakeslea trispora* entstanden genetisch veränderte Organismen. In *Blakeslea* liegen in allen Stadien des vegetativen und des sexuellen Zellzyklus mehrkernige Zellen vor. Daher erfolgte die Insertion der Vector-DNA in der Regel nur in einem Kern. Ziel ist es aber, dass die Insertion der Vector-DNA in allen Kernen vorliegt.



Zur Herstellung solcher homokaryotischer Zellen wurden zunächst Sporensuspensionen der rekombinanten Stämme mit MNNG behandelt. Hierfür wurde eine Sporensuspension mit  $1 \times 10^7$  Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100  $\mu\text{g/ml}$  zugegeben. Die Dauer der Inkubation mit MNNG wurde so gewählt, dass die Überlebensrate der Sporen  $\sim 5\%$  betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.

10

Anschließend wurden die Sporen auf MEP-Agarplatten ausplattiert und neue Sporen erzeugt.

15

Diese Sporen wurden analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin plattiert, das zusätzlich Hygromycin enthält.

Hierdurch wurden homokaryonte Zellen des Genotyps  $\text{hyg}^R$  und  $\text{dar}^-$  selektiert.

20

Nach diesem Prinzip wurden homokaryonte Stämme von *Blakeslea trispora* mit dem Phänotyp  $\text{hyg}^R$  und  $\text{dar}^-$  erzeugt.

Patentansprüche

1. Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* umfassend

5

(i) Transformation mindestens einer der Zellen,

(ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen durch Eliminierung von Kernen in diesen Zellen, so dass mindestens eine Zelle mit nur einem Kern verbleibt, der die gewünschte genetische Veränderung aus (i) aufweist und zur Ausprägung führt, und

10

(iii) Selektion der gentechnisch veränderten Zelle oder Zellen.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es sich um Zellen von Pilzen der Art *Blakeslea trispora* handelt.

3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector oder freie Nukleinsäuren verwendet werden.

15

4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector in das Genom mindestens einer der Zellen integriert wird.

20

5. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector derart gestaltet ist, dass die im Vector enthaltene genetische Information in der Zelle ausgeschaltet wird.

6. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector derart gestaltet ist, dass die im Vector enthaltene genetische Information in die Zelle eingeführt wird.

25

7. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 6, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector einen Promotor und/oder einen Terminator enthält.

8. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 7, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend den gpd Promotor und/oder den trpC Terminator eingesetzt wird.
- 5 9. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 8, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend ein Resistenzgen eingesetzt wird.
10. Verfahren nach Anspruch 8, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector ein Hygromycin-Resistenzgen (hph), insbesondere aus E. coli enthält.
- 10 11. Verfahren nach einem der vorhergehenden Ansprüche 7 - 10, **dadurch gekennzeichnet, dass** der gpd Promotor die Sequenz SEQ ID NO: 1 aufweist.
12. Verfahren nach einem der vorhergehenden Ansprüche 7 - 10, **dadurch gekennzeichnet, dass** der trpC Terminator die Sequenz SEQ ID NO: 2 aufweist.
- 15 13. Verfahren nach einem der Ansprüche 8 bis 12, **dadurch gekennzeichnet, dass** der gpd Promotor und der trpC Terminator aus Aspergillus nidulans stammen.
14. Verfahren nach einem Ansprüche 3 bis 13, **dadurch gekennzeichnet, dass** der Vector die SEQ ID NO: 3 umfasst.
- 20 15. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Transformation (i) mittels Agrobakterien, Konjugation, Chemikalien, Elektroporation, Beschuss mit DNA-beladenen Partikeln, Protoplasten oder Mikroinjektion durchgeführt wird.
16. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in der Homokaryontisierung (ii) ein mutagenes Agens eingesetzt wird.
- 25 17. Verfahren nach Anspruch 16, **dadurch gekennzeichnet, dass** als mutagenes Agens N-Methyl-N'-nitro-nitrosoguanidin (MNNG) eingesetzt wird.

18. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in der Selektion (iii) 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) eingesetzt werden.

5 19. Verfahren nach einem der Ansprüche 3 bis 18, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder deren Vorstufen enthält.

20. Verfahren nach einem der Ansprüche 3 bis 19, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinen oder Xanthophyllen enthält.

10 21. Verfahren nach einem der Ansprüche 3 bis 20, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon,  $\beta$ -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin,  $\beta$ -Carotin, Lutein oder  
15 Phytoen enthält.

22. Genetisch veränderte mehrkernige Zellen der zur Gattung Blakeslea gehörenden Pilze, insbesondere Blakeslea trispora erhältlich nach einem der vorhergehenden Ansprüche.

20 23. Verwendung der Zellen nach Anspruch 22 oder eines aus ihnen gebildeten Mycels zur Herstellung von Carotinoiden oder deren Vorstufen.

24. Verwendung nach Anspruch 22 oder 23 zur Herstellung von Carotinen oder Xanthophyllen.

25 25. Verwendung 22 bis 24 zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon,  $\beta$ -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin,  $\beta$ -Carotin, Lutein oder Phytoen.

26. Promotor mit der Sequenz SEQ ID NO: 1 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 21.

27. Terminator mit der Sequenz SEQ ID NO: 2 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 21.

28. Vector mit der SEQ ID NO: 3 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 21.

**Verfahren zur gentechnischen Veränderung von Organismen der Gattung**  
**Blakeslea, entsprechende Organismen und deren Verwendung**

**Zusammenfassung**

5

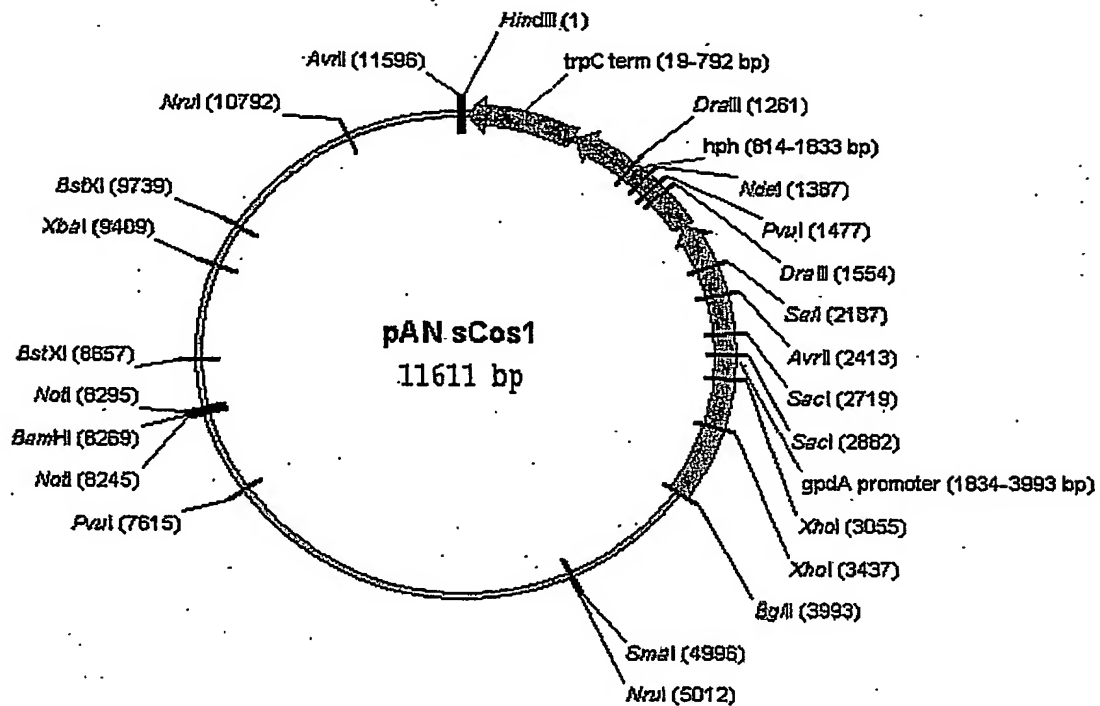
Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung Blakeslea umfassend

- (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryontisierung der aus (i) erhaltenen Zellen durch Eliminierung von Kernen in diesen Zellen, so dass mindestens eine Zelle mit nur einem Kern verbleibt, der die gewünschte genetische Veränderung aus (i) aufweist und zur Ausprägung führt, und
- (iii) Selektion der gentechnisch veränderten Zelle oder Zellen;

entsprechende Organismen und deren Verwendung.

15

**Fig. 1: Vektor pANsCos1**



**Fig. 2: Vektor pBinAHyg**

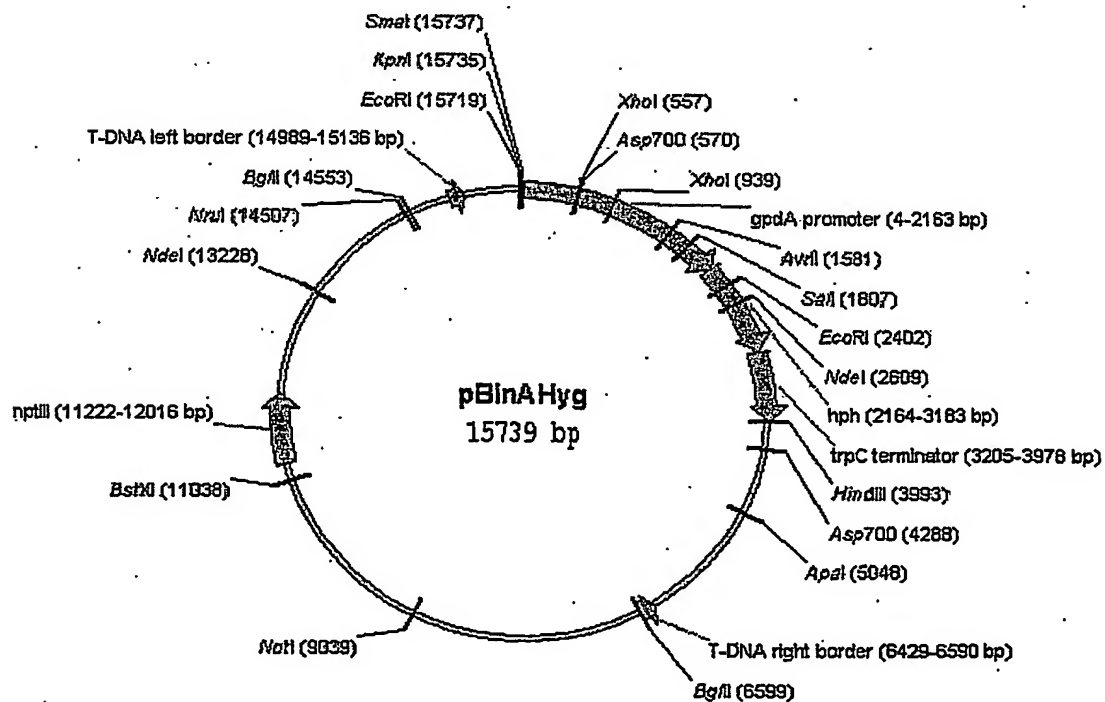
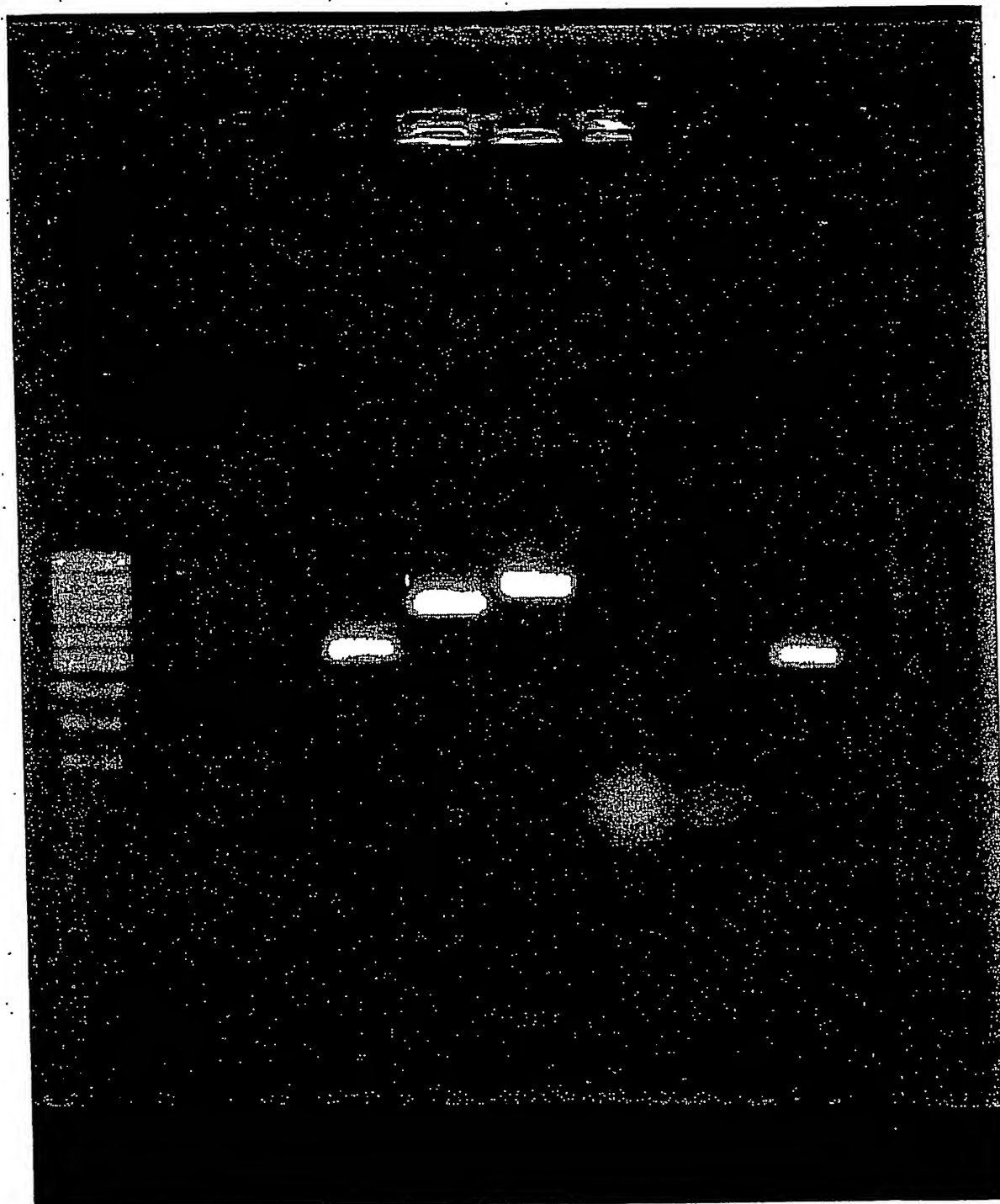


Fig. 3: Gels des Ergebnis einer PCR

Spur:

1 2 3 4 5 6 7 8 9



BEST AVAILABLE COPY



SEQUENZPROTOKOLL

BEST AVAILABLE COPY

<110> BASF AG

<120> Verfahren zur gezielten genetischen Veränderung von  
mehrkernigen Zel-len des Pilzes der Gattung Blakeslea,  
entsprechende Zellen und deren Verwendung

<130> BASF NAE 597/03

<140>

<141>

<160> 10

<170> PatentIn Ver. 2.1

<210> 1

<211> 2160

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Promotor

<400> 1

tttcgacac tgaatacgt cgagcctgct ccgcttggaa gcggcgagga gcctcgctct 60  
cacaacta ccaacatgga gtacgataag ggccagttcc gccagctcat taagagccag 120  
ttcatgggcg ttggcatgat ggccgtcatg catctgtact tcaagtacac caacgctctt 180  
ctgatccagt cgatcatccg ctgaaggcgc tttcgaatct ggttaagatc cacgtcttcg 240  
ggaagccagc gactggtgac ctccagcgtc cctttaaggc tgccaacagc tttctcagcc 300  
agggccagcc caagaccgac aaggcctccc tccagaacgc cgagaagaac tggaggggtg 360  
gtgtcaagga ggagtaagct ccttattgaa gtcggaggac ggagcgggtg caagaggata 420  
ttcttcgact ctgtattata gataagatga tgaggaattg gaggtagcat agcttcattt 480  
ggatttgctt tccaggctga gactctagct tggagcatag agggtccttt ggctttcaat 540  
attctcaagt atctcgagtt tgaacttatt cctgtgaac cttttattca ccaatgagca 600  
ttggaatgaa catgaatctg aggactgcaa tcgccatgag gttttcgaaa tacatccgga 660  
tgtcgaaggc ttggggcacc tgcgttggtt gaatttagaa cgtggcacta ttgatcatcc 720  
gatagctctg caaagggcgt tgcacaatgc aagtcaaagc ttgctagcag ttccagggtg 780  
aatgttatga tgagcattgt attaaatcag gagatatagc atgatctcta gttagctcac 840  
cacaaaagtc agacggcgta accaaaagtc acacaacaca agctgtaagg atttcggcac 900

ggctacggaa gacggagaag ccaccttcag tggactcgag taccatttaa ttctatttgt 960  
gtttgatcga gacctaatac agccccctaca acgaccatca aagtcgtata gctaccagtg 1020  
aggaagtgga ctcaaatega cttcagcaac atctcctgga taaactttaa gcctaaacta 1080  
tacagaataa gatagggtgga gagcttatac cgagctccca aatctgtcca gatcatgggt 1140  
gaccggtgcc tggatcttcc tatagaatca tccttattcg ttgacctagc tgattctgga 1200  
gtgaccacga gggatcatgac ttgagcctaa aatccgcccgc ctccaccatt tgtagaaaaa 1260  
tgtgacgaac tcgtgagctc tgtacagtga ccggtgactc tttctggcat gcggagagac 1320  
ggacggacgc agagagaagg gctgagtaat aagccactgg ccagacagct ctggcgggctc 1380  
tgagggtgcag tggatgatta ttaatccggg accggcggcc cctccgcccc gaagtggaaa 1440  
ggctgggtgtg cccctcgttg accaagaatc tattgcatca tcggagaata tggagcttca 1500  
tcgaatcacc ggcagtaagc gaaggagaat gtgaagccag ggggtgtatag ccgtcggcga 1560  
aatagcatgc cattaacctg ggtacagaag tccaattgct tccgatctgg taaaagattc 1620  
acgagatagt accttctccg aagtaggtag agcgagtacc cggcgcgtaa gctccctaatt 1680  
ggcccatcc ggcatctgta gggcgtccaa atatcgtgcc tctcctgctt tgcccgggtgt 1740  
gaaaccgg aaaggccgct caggagctgg ccagcggcgc agaccgggaa cacaagctgg 1800  
cagtcgaccc atccggtgct ctgcactcga cctgctgagg tccctcagtc cctggtaggc 1860  
agctttgccc cgtctgtccg ccggtgtgt cggcgggggt gacaaggctc ttgcgtcagt 1920  
ccaacatttg ttgccatatt ttcctgctct cccaccagc tgctcttttc ttttctcttt 1980  
cttttcccat cttcagtata ttcattcttc catccaagaa cctttatttc ccctaagtaa 2040  
gtactttgct acatccatac tccatccttc ccatccctta ttcctttgaa cttttcagtt 2100  
cgagctttcc cacttcatcg cagcttgact aacagctacc ccgcttgagc agacatcacc 2160

<210> 2

<211> 774

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Terminator

<400> 2

cgatccactt aacgttactg aaatcatcaa acagcttgac gaatctggat ataagatcgt 60  
tgggtgctgat gtcagctccg gagttgagac aaatggtgtt caggatctcg ataagatacg 120  
ttcatattgtc caagcagcaa agagtgcctt ctagtgattt aatagctcca tgtcaacaag 180  
aataaaaacgc gttttcgggt ttacctcttc cagatacagc tcatctgcaa tgcattaatg 240  
cattgactgc aacctagtaa cgccttncag gctccggcga agagaagaat agcttagcag 300  
agctatttttc attttcggga gacgagatca agcagatcaa cggtcgtcaa gagacctacg 360  
agactgagga atccgctctt ggctccacgc gactatatat ttgtctctaa ttgtactttg 420  
acatgctcct cttctttact ctgatagctt gactatgaaa attccgtcac cagcncctgg 480  
gttcgcaaag ataattgcat gtttcttctc tgaactctca agcctacagg acacacattc 540

atcgtaggta taaacctcga aatcanttcc tactaagatg gtatacaata gtaaccatgc 600  
atggttgccct agtgaatgct ccgtaacacc caatacgccg gccgaaactt ttttacaact 660  
ctcctatgag tcgtttaccc agaatgcaca ggtacacttg ttttagaggta atccttcttt 720  
ctagctagaa gtcctcgtgt actgtgtaag cgcccactcc acatctccac tcga 774

<210> 3

<211> 15739

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Vector

00> 3

atcttttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60  
cctgtcacaa ctaccaacat ggagtagcat aagggccagt tccgccagct cattaagagc 120  
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180  
cttctgatcc agtcgatcat ccgctgaagg cgcttttcgaa tctgggtaag atccacgtct 240  
tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca 300  
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360  
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420  
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480  
tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540  
aatattctca agtatctcga gtttgaactt attccctgtg aaacttttat tcaccaatga 600  
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatcacatcc 660  
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720  
cgatagct ctgcaaaggc cgcttcacaa tgcaagtcaa acgttgctag cagttccagg 780  
ggaatgta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840  
caccacaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg 900  
cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960  
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca 1020  
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080  
ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140  
gttgaccggt gcctggatct tcctatagaa tcatccttat tcgttgacct agctgattct 1200  
  
ggagtgaccc agaggggtcat gacttgagcc taaaatccgc cgctccacc atttgtagaa 1260  
aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320  
gacggacgga cgagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380  
ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440  
aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560  
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgac tggtaaaaga 1620  
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680  
aattggccca tccggcatct gtagggcgctc caaatatcgt gcctctcctg ctttgcccgg 1740  
tgtatgaaac cggaaaggcc gtcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800  
tggcagtcga cccatccggg gctctgcact cgacctgctg aggtccctca gtccctggta 1860  
ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgctc 1920  
agtccaacat ttgttgccat attttctctg tctccccacc agctgctctt ttcttttctc 1980  
tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040  
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100  
gttcgagctt tccacttca tcgcagcttg actaacagct acccgccttg agcagacatc 2160  
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220  
gtctccgacc tgatgcagct ctccggaggc gaagaatctc gtgctttcag cttcgatgta 2280  
tggggcggtg gatatgtcct gcgggtaaat agctgcgcgc atggtttcta caaagatcgt 2340  
tggtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400  
gaattcagcg agagcctgac ctattgcac tcccgcctg cacaggggtg cacgttgcaa 2460  
gacctgcctg aaaccgaact gccgcgtgtt ctgcagccgg tcgcggaggc catggatgcg 2520  
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580  
ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640  
tggcaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700  
atgctttggg ccgaggactg ccccgaaagtc cggcacctcg tgcacgcgga tttcggctcc 2760  
aacaatgtcc tgacggacaa tggccgcata acagcggta ttgactggag cgaggcgatg 2820  
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880  
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgacgg atcgccgcgg 2940  
ctccgggcgt atatgctccg cattggctct gaccaactct atcagagctt ggttgacggc 3000  
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctcc atccggagcc 3060  
tgactgtcg ggctacaca aatcgccgcg agaagcgcg ccgtctggac cgatggctgt 3120  
tgaagtac tcgccgatag tggaaaccga cgcgccagca ctgctccgag ggcaaaggaa 3180  
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240  
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300  
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagt ccttctagt 3360  
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420  
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480  
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540  
tcaacggctg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600  
atatttgtct ctaattgtac ttgacatgc tctcttctt tactctgata gcttgactat 3660  
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tcttgaaact 3720  
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780  
gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa caccaatac 3840  
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900  
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caagcttggc gtaatcatgg tcatagctgt 4020  
ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 4080  
agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg ttgctgtcac 4140  
tgccccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 4200  
cggggagagg cggtttgctt attggggcaa agacaaaagg gcgacattca accgattgag 4260  
ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4320  
cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4380  
ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4440  
caagtttgcc tttagcgtca gactgtagcg cgttttcacg ggcattttcg gtcatagccc 4500  
ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg 4560  
aaccgcctcc ctcagagccg ccaccctcag aaccgccacc ctcagagcca ccaccctcag 4620  
agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt 4680  
aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4740  
atcgctgatt aaatgtataa ttgctgggact ctaatcataa aaacccatct cataaataac 4800  
atcatcgcaa gaccggcaac aggattcaat cttaagaaac tttattgcca aatgtttgaa 4920  
cgatcgggga tcatccgggt ctgtggcggt aactccacga aaatatccga acgcagcaag 4980  
atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg 5040  
ccgggcccga tcatattgtc gctcaggatc gtggcggtgt gcttgtcggc cgttgctgtc 5100  
gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccggtggc gaagaactcc 5160  
agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag 5220  
cccaaccttt catagaaggc ggcgggtgaa tcgaaatctc gtgatggcag gttgggctgc 5280  
gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat 5340  
agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5400  
cccattcgcc gccaaactct tcagcaatat cacgggtagc caacgctatg tcctgatagc 5460  
ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca ttttccacca 5520  
tgatattcgg caagcaggca tcgcatggg tcacgacgag atcatcgccg tcgggcatgc 5580  
gccttgag cctggcgaac agttcggtg gcgcgagccc ctgatgctct tcgtccagat 5640  
atcctgatc gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5700  
cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag 5760  
ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca 5820  
cttcgcccga tagcagocag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc 5880  
aaggaaacgc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc agttcattca 5940  
gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg ccctgctgct gacagccgga 6000  
acacggcggc atcagagcag ccgattgtct gttgtgcccga gtcatagccg aatagcctct 6060  
ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgacg 6120  
cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 6180  
gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6240  
accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6300  
actccagaaa cccgcggctg agtggctect tcaacgttgc ggttctgtca gttccaaacg 6360  
taaaacggct tgtcccgct catcgccggg ggtcataacg tgactccctt aattctccgc 6420

tcattgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg acaggatata 6480  
ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat ttaaaagggc 6540  
gtgaaaagggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga 6600  
tctggcgccg gccagcgaga cgagcaagat tggccgcccgc ccgaaacgat ccgacagcgc 6660  
gcccagcaca ggtgocgagg caaattgcac caacgcatac agcgccagca gaatgccata 6720  
gtgggcgggtg acgtcgttcg agtgaaccag atcgcgccagg agggcccggca gcaccggcat 6780  
aatcaggccg atgccgacag cgtcgagcgc gacagtgttc agaattacga tcaggggtat 6840  
gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggatt 6900  
ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6960  
acaaagttgc agccgaatac agtgcacgtt gccgccttgg acctgttgaa cgaggtcggc 7020  
gtagacgggtc tgacgacacg caaactggcg gaacggttgg ggggttcagca gccggcgctt 7080  
tactggcact tcaggaacaa gcggggcgctg ctcgacgcac tggccgaagc catgctggcg 7140  
gagaatcata cgcattcggt gccgagagcc gacgacgact ggcgctcatt tctgatcggg 7200  
tgcccgcga gcttcaggca ggcgctgttc gcctaccgcg atggcgcgcg catccatgcc 7260  
gcacgcgac cgggcgaccc gcagatggaa acggccgacg cgcagcttcg ctctctctgc 7320  
gaggcgggtt tttcgcccg ggacgccgtc aatgcgctga tgacaatcag ctacttcaact 7380  
gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7440  
gttgaacagg ctccgctctc gccgctgttg cgggccgca tagacgcctt cgacgaagcc 7500  
ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7560  
aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7620  
ccggagcgca acccactcac tacagcagag ccatgtagac aacatccctt ccccttttcc 7680  
accgcgtcag acgcccgtag cagcccgtca cgggcttttt catgccctgc cctagcgtcc 7740  
aagcctcacg gccgcgtcg gcctctcttg cggccttctg gcgctcttcc gcttctctgc 7800  
tactgactc gctgcgtcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 7860  
cggtaatacgt gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 7920  
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttggc ggcgtttttc cataggctcc 7980  
ccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag 8040  
actataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 8100  
ccctgcgct taccggatac ctgtccgctt ttctccctt gggagcgtg gcgcttttcc 8160  
gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc atcctttttc 8220  
gcacgatata caggattttg ccaaagggtt cgtgtagact ttccttggtg tatccaacgg 8280  
cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc cttcttcaact 8340  
gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccgggt 8400  
accgcccggc taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8460  
ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8520  
gcggcgggcg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa 8580  
atcacggggc tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa tggcgacctg 8640  
ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgac ggcgcggttc 8700  
ggtgatgcc aatcctctgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8760  
aaggtcatga tggcggtggt ccgcccagg gcagagccat gactttttta gccgctaaaa 8820

BEST AVAILABLE COPY

cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8880

cttcgcggag ctggtgaagt acatcaccca cgagcaaggc aagaccgagc gcctttgcga 8940

cgctcacccg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc cctgcaaacc 9000

cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt tgtggatacc 9060

tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgagggggccg 9120

actcacccgg cgcggcgctt acagatgagg ggcaggctcg atttcggccg gcgacgtgga 9180

gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc ccacagatga 9240

tgtggacaag cctgggggata agtgccctgc ggtattgaca cttgagggggc gcgactactg 9300

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgagggggcgc 9360

acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc aagggtttcc 9420

gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9480

acctgttttt taaccagggc tgcgcctgtg gcgcgtgacc gcgcacgccg aaggggggtg 9540

cccccccttc tcgaacctc ccggcccgct aacgcgggccc tcccatcccc ccagggggctg 9600

ccctcggc cgcgcaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9660

gccgggatcg ggcagtaac gggatgggcg atcagcccga gcgcgacgcc cggaagcatt 9720

gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9780

ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga cttcatggcg 9840

gggccggcaa tttttacctt gggcattctt ggcatagttg tcgcgggtgc cgtgctcgtg 9900

ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggttaagatt ataccgaggt 9950

atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat ttaaaaagct 10020

accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 10080

ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 10140

caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 10200

catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10260

tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc gatgactttg 10320

tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10380

ccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc 10440

agctttccct tcaggcgga ttcatacagc ggccagccat ccgtcatcca tatcaccacg 10500

tcaaagggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg ttcaccgaat 10560

acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca gcgctggcgc 10620

gatttagccc cgacatagcc ccactgttcg tccatttccg cgcagacgat gacgtcactg 10680

cccggtgta tgccgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10740

tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca ttcattggcca 10800

tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac tgcagttgcc 10860

atgttttacg gcagtgaag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10920

gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10980

cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 11040

gtttcaaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 11100

aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gaggttcgtct 11160

tggtataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 11220

BEST AVAILABLE COPY

aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11280  
aaaagatacg gaaggaatgt ctcctgctaa ggtatataag ctggtgggag aaaatgaaaa 11340  
cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11400  
aaaggacatg atgctatggc tggaaggaaa gctgcctggt ccaaagggtcc tgcactttga 11460  
acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgctcc tttgctcgga 11520  
agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcacag 11580  
gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11640  
agccgaattg gattacttac tgaataacga tctggccgat gtggattgcy aaaactggga 11700  
agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11760  
cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct ttgtgaaaga 11820  
tggaagaa agtggtctta ttgatcttgg gagaagcggc agggcggaaga agtggtatga 11880  
cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct 11940  
attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 12000  
atgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 12060  
gacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc aagtatttgg 12120  
gcaaggggtc gctggtattc gtgcagggca agattcgga taccaagtac gagaaggacg 12180  
gccagacggg ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12240  
caccaggcgg gtcaaactag gaataagggc acattgcccc ggctgagtc ggggcaatcc 12300  
cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12360  
acgcgggggt ttccgccgag gatgccgaaa ccacgcgaag ccgcaccgtc atgcgtgcgc 12420  
cccgcgaaac cttccagtc gtcggctcga tggccagca agctacggcc aagatcgagc 12480  
gcgacagcgt gcaactggct cccctgccc tggccgcgc atcgccgcgc gtggagcggt 12540  
cgctcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc gacacgcgag 12600  
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12660  
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12720  
ccttgctcga tattgcgcgc tggccggaca cgatgcgagc gatgcaaac gacacggccc 12780  
ctctgcct gttcaccac cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12840  
attttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag ctgcgggccg 12900  
acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc cctatcggcg 12960  
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg atcaatggcc 13020  
ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 13080  
cgtccgaccg cgttgggcac ctggaatcgg tgctcgtgct gcaccgcttc cgcgtcctgg 13140  
accgtggcaa gaaaacgtcc cgttggcagg tcctgatcga cgaggaaatc gtcgtgctgt 13200  
ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13260  
cccagcgat gttegactat ttcagctcgc accgggagcc gtaccgcctc aagctggaaa 13320  
ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gacaggtcg 13380  
gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg 13440  
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg gggttcagcag 13500  
ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13560  
agtatcgctc gggacgcacg gcgcgtcta cgaactgccg ataaacagag gattaaaatt 13620  
gaatattgtg attaaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13680

BEST AVAILABLE COPY



cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg tttacgagca 13740  
cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13800  
cgcctacatc gacggcgaga tcattgggct gtcgggtcttc aaacaggagg acggccccaa 13860  
ggacgctcac aaggcgcac tcgtccggcg tttcgtggag cccgaacagc gaggccgagg 13920  
ggtcgcgggt atgctgctgc gggcggttgc ggcgggttta ttgctcgtga tgatcgtccg 13980  
acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac ttaatatctc 14040  
gctattctgg agcttggtgt ttatttcggg ctaccgcctg cggggcgggg tcgcgggcag 14100  
ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc taggtagccc 14160  
gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggt 14220  
gttgacacca aacgcagcgc tagatcctgt cggcgctcga gcgggcctgg cgggggagggt 14280  
ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgtgc ctctgctcac 14340  
ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag ctttagtggt 14400  
tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcggt ggctcggcct 14460  
tcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14520  
ttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cggggatgca 14580  
tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag caatggatag 14640  
gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc ttcctcagcg 14700  
gctttatcca gcgatttcct attatgtcgg catagttctc aagatcgaca gcctgtcacg 14760  
gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg agatgatatt 14820  
tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14880  
atccgtgttt caaaccgggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14940  
gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 15000  
ctgtatcgag tggtgatttt gtgcccagct gccggtcggg gagctgttgg ctggctgggtg 15060  
gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 15120  
cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa cagctgattg 15180  
cccttcacgc cctggccctg agagagttgc agcaagcggg ccacgctggt ttgccccagc 15240  
tcgaaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15300  
ctagcccg gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga 15360  
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15420  
aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta aatcgggaacc 15480  
ctaaagggag ccccgatttt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 15540  
aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15600  
tcgggtcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 15660  
ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac ggccagtga 15720  
ttcgagctcg gtacccggg 15739

<210> 4

<211> 11611

<212> DNA

<213> Künstliche Sequenz

BEST AVAILABLE COPY

<220>

<223> Beschreibung der künstlichen Sequenz:Vector

<400> 4

agcttgcatg cctgcaggtc gaggaggat gtggagtggt cgcttacaca gtacacgagg 60  
acttctagct agaaagaagg attacctcta aacaagtgtta cctgtgcatt ctgggtaaac 120  
gactcatagg agagttgtta aaaagtttcg gccggcgat tgggtgttac ggagcattca 180  
ctaggcaacc atgcatgggt actattgtat accatcttag taggaantga tttcgagggt 240  
tatacctacg atgaatgtgt gtctgttagg cttgagagtt caaggaagaa acatgcaatt 300  
atctttgcga acccaggngc tgggtgacgga attttcatag tcaagctatc agagtaaaga 360  
agaggagcat gtcaaagtac aattagagac aaatatatag tcgctggag ccaagagcgg 420  
attcctcagt ctcttaggtc tcttgacgac cgttgatctg cttgatctcg tctccgaaa 480  
gaaaatag ctctgctaag ctattcttct cttcgccgga gcctgmaagg cgttactagg 540  
gcagtcaa tgcattaatg cattgcagat gagctgtatc tggaagaggt aaacccgaaa 600  
acgcgtttta ttcttggtga catggagcta ttaaatcact agaaggcact ctttgctgct 660  
tggaacaaatg aacgtatctt atcgagatcc tgaacaccat ttgtctcaac tccggagctg 720  
acatcgacac caacgatctt atatccagat tcgtcaagct gtttgatgat ttcagtaacg 780  
ttaagtggat cgatcccgcg gtcggcatct actctattcc tttgccctcg gacgagtgt 840  
ggggcgctcg tttccactat cggcgagtag ttctacacag ccacggtcc agacggccgc 900  
gcttctgcgg gcgatttgtg tacgcccgc agtcccggct ccggtatcga cgattgcgtc 960  
gcacgcaccc tgcgcccaag ctgcacatc gaaattgccg tcaaccaagc tctgatagag 1020  
ttggtcaaga ccaatgcgga gcatatacgc ccggagccgc ggcgatcctg caagctccgg 1080  
atgcctccgc tcgaagtagc gcgtctgtg ctccatacaa gccaacacg gcctccagaa 1140  
gaagatgttg gcgacctcgt attgggaatc cccgaacatc gcctcgctcc agtcaatgac 1200  
cgctgttatg cggccattgt ccgtcaggac attgttgag ccgaaatccg cgtgcacgag 1260  
tgccggact tcggggcagt cctcggccca aagcatcagc tcacgagag cctgcgcgac 1320  
gacgcactg acggtgtcgt ccacacagt ttgccagtga tacacatggg gatcagcaat 1380  
cgcgcatatg aaatcacgcc atgtagtgt ttgaccgatt ccttgcggtc cgaatgggccc 1440  
gaacccgctc gtctggctaa gatcggccgc agcgatcgca tccatggcct ccgcgaccgg 1500  
ctgcagaaca gcgggcagtt cggtttcagg caggctctgc aacgtgacac cctgtgcacg 1560  
gcgggagatg caatagggtca ggctctcgt gaattcccca atgtcaagca cttccggaat 1620  
cgggagcgcg gccgatgcaa agtgccgata aacataacga tctttgtaga aaccatcggc 1680  
gcagctatctt acccgagga catatccacg ccctcctaca tcgaagctga aagcacgaga 1740  
ttcttcgccc tccgagagct gcatcaggtc ggagacgctg tcgaactttt cgatcagaaa 1800  
cttctcgaca gacgtcgcg tgagttcagg catggtgatg tctgtcaag cggggtagct 1860  
gttagtcaag ctgcgatgaa gtgggaaagc tcgaactgaa aggttcaaag gaataagggg 1920  
tggaagagat ggagtatgga ttagtcaaaag tacttactta ggggaaataa aggttcttgg 1980  
atgggaagat gaatatactg aagatgggaa aagaaagaga aaagaaaaga gcagctggtg 2040  
gggagagcag gaaaatatgg caacaaatgt tggactgacg caacgacctt gtcaaccccc 2100  
ccgacacacc gggcgagacg acggggcaaa gctgcctacc agggactgag ggacctcagc 2160

BEST AVAILABLE COPY

aggtcgagtg cagagcaccg gatgggtoga ctgccagctt gtgttcccgg tctgcgccgc 2220  
tggccagctc ctgagcgggc ttcccggttt catacaccgg gcaaagcagg agaggcacga 2280  
tatttgagcg ccetacagat gccggatggg ccaattaggg agcttacgcg ccgggtactc 2340  
gctctaccta cttcggagaa ggtactatct cgtgaatctt ttaccagatc ggaagcaatt 2400  
ggacttctgt acctaggtta atggcatgct atttcgccga cggctatata cccctggctt 2460  
cacattctcc ttcgcttact gccggtgatt cgatgaagct ccatattctc cgatgatgca 2520  
atagattctt ggtcaacgag gggcacacca gcctttccac ttcggggcgg aggggcggcc 2580  
ggtcccggat taataatcat ccaactgcacc tcagagccgc cagagctgtc tggccagtgg 2640  
cttattactc agcccttctc tctgcgtccg tccgtctctc cgcatgccag aaagagtcac 2700  
cggtcactgt acagagctca cgagttcgtc acatttttct acaaagggtg gaggcggcgg 2760  
attttaggct caagtcatga cccctctgggt cactccagaa tcagctaggt caacgaataa 2820  
ggatgattct ataggaagat ccaggcaccg gtcaaccatg atctggacag atttgggagc 2880  
tcggtataag ctctccacct atcttattct gtatagttta ggcttaaagt ttatccagga 2940  
tggttgctg aagtcgattt gagtccactt cctcactggt agctatacga ctttgatggt 3000  
ttgtaggg gctgtattag gtctcgatca aacacaaata gaattaaatg gtactcgagt 3060  
ccactgaagg tggcttctcc gtcttcogta gccgtgccga aatccttaca gcttggtgtg 3120  
tgtgactttt ggttacgccg tctgactttt gtggtgagct aactagagat catgctatat 3180  
ctcctgattt aatacaatgc tcatcataac attccacctg gaactgctag caacgtttga 3240  
cttgcatgtt gcaacgccct ttgcagagct atcggatgat caatagtgcc acgttctaaa 3300  
ttcaaccaac gcaggtgccc caagccttcg acatccggat gtatttcgaa aacctcatgg 3360  
cgattgcagt cctcagattc atgttcattc caatgctcat tggatgaataa aaggttcaca 3420  
gggaataagt tcaaaactga gatacttgag aatattgaaa gccaaaggac cctctatgct 3480  
ccaagctaga gtctcagcct ggaaagcaaa tccaaatgaa gctatgtac ctccaattcc 3540  
tcatcatctt atctataata cagagtcgaa gaatatctc ttgacaccgc tccgtcctcc 3600  
gacttcaata aggagcttac tcctccttga caccaccct ccagttcttc tcggcggtct 3660  
ggagggaggc cttgtcggtc ttgggctggc cctggctgag aaagctggtg gcagccttaa 3720  
gggacgctg gaggtcacca gtcgctggct tcccgaagac gtggatctta accagattcg 3780  
aagcgcctt cagcggatga tcgactggat cagaagagcg ttggtgtact tgaagtacag 3840  
atgcatgacg gccatcatgc caacgcccac gaactggctc ttaatgagct ggcggaactg 3900  
gcccttatcg tactccatgt tggtagttgt gacaggacga ggctcctcgc cgcttccaag 3960  
cggagcaggc tcgacgtatt tcagtgtcga aagatctgat caagagacag gatgaggatc 4020  
gttttcgcatg attgaacaag atggattgca cgcaggttct ccggccgctt gggtaggag 4080  
gctattcggc tatgactggg cacaacagac aatcggctgc tctgatgccg ccgtgttccg 4140  
gctgtcagcg caggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa 4200  
tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacgggcg ttccttgcc 4260  
agctgtgtc gacgttgtca ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc 4320  
ggggcaggat ctctgtcat ctcaccttgc tcctgccgag aaagtatcca tcatggctga 4380  
tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 4440  
acatcgcac gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct 4500  
ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat 4560  
gcccgcggc gaggatctcg tcgtgaccca tggcgatgcc tgcttgccga atatcatggt 4620

ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 4680  
tcaggacata gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga 4740  
ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg 4800  
ccttcttgac gagttcttct gagegggact ctggggttcg aaatgaccga ccaagcgacg 4860  
cccaacctgc catcacgaga tttcgattcc accgccgcct tctatgaaag gttgggcttc 4920  
ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct catgctggag 4980  
ttcttcgccc accccgggct cgatcccctc gcgagttggt tcagctgctg cctgaggctg 5040  
gacgacctcg cggagtctta ccggcagtgc aaatccgtcg gcatccagga aaccagcagc 5100  
ggctatccgc gcatccatgc ccccgaactg caggagtggg gaggcacgat ggccgctttg 5160  
gtccggatct ttgtgaagga accttacttc tgtggtgtga cataattgga caaactacct 5220  
acagagattt aaagctctaa ggtaaataa aaatttttaa gtgtataatg tgttaaacta 5280  
ctgattctaa ttgtttgtgt atttttagatt ccaacctatg gaactgatga atgggagcag 5340  
tggtggaatg cctttaatga ggaaaacctg ttttgctcag aagaaatgcc atctagtgat 5400  
gatgaggcta ctgctgactc tcaacattct actcctccaa aaaagaagag aaaggtagaa 5460  
ccccaagg actttccttc agaattgcta agttttttga gtcatgctgt gtttagtaat 5520  
agaactcttg cttgctttgc tatttacacc acaaaggaaa aagctgcact gctatacaag 5580  
aaaattatgg aaaaatattc tgtaaccttt ataagtaggc ataacagtta taatcataac 5640  
atactgtttt ttcttactcc acacaggcat agagtgtctg ctattaataa ctatgctcaa 5700  
aaattgtgta ccttttagctt ttttaatttg aaaggggtta ataaggaata tttgatgtat 5760  
agtgccttga ctagagatca taatcagcca taccacattt gtagaggttt tacttgcttt 5820  
aaaaaacctc ccacacctcc ccctgaacct gaaacataaa atgaatgcaa ttgttggtgt 5880  
taacttgttt attgcagctt ataatgggta caaataaagc aatagcatca caaatttcac 5940  
aaataaagca tttttttcac tgcattctag ttgtggttg tccaaactca tcaatgtatc 6000  
ttatcatgtc tggatctgac ggggtgcgcat gatcgtgctc ctgtcgttga ggacccggct 6060  
aggctggcgg gggtgcctta ctggttagca gaatgaatca ccgatacgcg agcgaacgtg 6120  
aagcgactgc tgctgcaaaa cgtctgcgac ctgagcaaca acatgaatgg tcttcggttt 6180  
ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgctc ttccgcttcc tcgctcactg 6240  
tcgctgcg ctcggtcggt cggtgcggc gagcggatc agctcactca aaggcggtaa 6300  
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 6360  
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 6420  
gctccgcccc cctgacgagc atcacaaaaa tgcagctca agtcagaggt ggcgaaacct 6480  
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 6540  
tccgacctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 6600  
ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 6660  
ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct 6720  
tgagtccaac ccggaagac acgacttate gccactggca gcagccactg gtaacaggat 6780  
tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtggc ctaactacgg 6840  
ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 6900  
aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcgggt gtttttttgt 6960  
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 7020  
tacgggggtct gacgctcagt ggaacgaaaa ctacagttaa gggattttgg tcatgagatt 7080

BEST AVAILABLE COPY

atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 7140  
aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 7200  
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac 7260  
tacgatacgg gagggcttac catctggccc cagtgcctga atgataccgc gagaccacag 7320  
ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaa 7380  
tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 7440  
aagtagttcg ccagttaata gtttgcgcaa cgttggtgce attgctgcag gcatcgtggt 7500  
gtcacgctcg tegtgttgta tggcttcatt cagctccggt tcccaacgat caaggcgagt 7560  
tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggctctc cgtatcgttgt 7620  
cagaagtaag ttggccgcag tggtatcact catggttatg gcagcactgc ataattctct 7680  
tactgtcatg ccacccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt 7740  
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac 7800  
cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa 7860  
gtctcaagg atcttacgc tgttgagatc cagttcgatg taaccactc gtgcacccaa 7920  
gatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7980  
aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct 8040  
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 8100  
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 8160  
tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag 8220  
gccctttcgt cttcaagaat tcgcggccgc aattaaccct cactaaagga tccctatagt 8280  
gagtcgtatt atgcggccgc gaattctcat gtttgaccgc ttatcatcga taagctctgc 8340  
tttttgttga cttccattgt tcattccacg gacaaaaaca gagaaaggaa acgacagagg 8400  
ccaaaaagct cgctttcagc acctgtcgtt tcctttcttt tcagagggta ttttaaataa 8460  
aaacattaag ttatgacgaa gaagaacgga aacgccttaa accggaaaat tttcataaat 8520  
agcgaaaacc cgcgaggtcg ccgccccgta acaaggcgga tcgcccggaa ggaccgcgaa 8580  
atgataataa ttatcaattg catactatcg acggcactgc tgccagataa caccaccggg 8640  
aaacattcc atcatgatg ccgtgcccgc ataggaagcc agttcatcca tcgctttctt 8700  
ctgctgcc atttgctttg tgacatccag cgccgcacat tcagcagcgt ttttcagcgc 8760  
gttttcgatc aacgtttcaa tgttggtatc aacaccaggt ttaactttga acttatcggc 8820  
actgacgggt acctgtttct gcgctggctc atcacgcagg ataccaaggc tgatgttgta 8880  
gatattggtc accggctgag ggttttcgat tgccgctgcg tggatagcac catttgcgat 8940  
caggcngtcc ttgatgaatg aactccatt gcgaataagt tcgaaggaga cgggtgcacg 9000  
aatgcgctgg tccagctcgg tcgattgcct tttgtgcagc agaggtatca atctcaacgc 9060  
caaggctcat cgaagcgcaa tattgctgct caccaaaacg cgtattgacc aggtgttcaa 9120  
cggcaaattt ctgcccttct gatgtcagaa aggcaaagtg attttctttc tggatttcag 9180  
ttgtgtgtg tcggtttcag caaaaccaag ctgcgcgaat tcggctgtgc agatttagaa 9240  
ggcagatcac cagacagcaa cggccaacgg aaaacagcgc atacagaaca tccgtcgccg 9300  
cgccgacaac gtgataattt ttatgaccca tgatttatct ccttttagac gtgagcctgt 9360  
cgacagcaa agccgccgaa agttcctcga agctagcttc agacgtgtct agatacgtct 9420  
gctttttgtt gacttccatt gttcattcca cggacaaaaa cagagaaagg aaacgacaga 9480  
ggccaaaaag ctgcctttca gcacctgtcg tttcctttct tttcagaggg tatttttaaat 9540

```

aaaaacatta agttatgacg aagaagaacg gaaacgcctt aaaccggaaa attttcataa 9600
atagcgaaaaa cccgcgaggt cgccgccccg taacaaggcg gatcgccgga aaggacccgc 9660
aaatgataat aattatcaat tgcatactat cgacggcact gctgccagat aacaccaccg 9720
gggaaacatt ccattcatgat ggccgtgctg acataggaag ccagttcatc catcgctttc 9780
ttgtctgctg ccatttgctt tgtgacatcc agcgccgcac attcagcagc gtttttcagc 9840
gcgttttcga tcaacgtttc aatgttggtt tcaacaccag gtttaacttt gaacttatcg 9900
gcactgacgg ttaccttggt ctgcgctggc tcatcacgca ggataccaag gctgatgttg 9960
tagatattgg tcaccggctg aggggttttcg attgcccgtg cgtggatagc accatttgcg 10020
atcaggcngt ccttgatgaa tgacactcca ttgcgaataa gttcgaagga gacgggtgtca 10080
cgaatgcgct ggtccagctc ggtcgattgc cttttgtgca gcagaggtat caatctcaac 10140
gccaaaggctc atcgaagcgc aatattgctg ctcacaaaaa cgcgattga ccagggtgtc 10200
aacggcaaat ttctgccctt ctgatgtcag aaaggcaaag tgattttctt tctggtattc 10260
agttgctgtg tgtcggtttc agcaaaacca agctcgcgca attcggctgt gcagatttag 10320
aaggcagatc accagacagc aacggccaac ggaaaacagc gcatacagaa catccgtcgc 10380
cgccgaca acgtgataat ttttatgacc catgatttat ttccttttag acgtgagcct 10440
gtcgcacagc aaagccgccc aaagtctctc gaccgatgcc cttgagagcc ttcaaccag 10500
tcagctcctt ccggtgggcg cggggcatga ctatcgctgc cgacttatg actgtcttct 10560
ttatcatgca actcgtagga caggtgcccg cagcgtctct ggtcattttc ggcgaggacc 10620
gctttcgtg gagcgcgacg atgatcgcc tgctgcttgc ggtattcgga atcttgacg 10680
ccctcgctca agccttcgtc actgggtccc ccaccaaacy ttcggcgag aagcaggcca 10740
ttatcgccgg catggcgccc gacgcgtgg gctacgtctt gctggcgctc gcgacgcgag 10800
gctggatggc cttccccatt atgattcttc tcgcttcggc cggcatcggg atgcccgcgt 10860
tgcaggccat gctgtccagg caggtagatg acgaccatca gggacagctt caaggatcgc 10920
tcgcggtctt taccagccta acttcgatca ttggaccgct gatcgtcacg gcgatttatg 10980
ccgcctcggc gagcacatgg aacgggttgg catggattgt aggcgcgcgc ctataccttg 11040
tctgcctccc cgcgttgct cgcggtgcat ggagccgggc cacctcgacc tgaatggaag 11100
ccggcgccac ctgcctaacy gattcaccac tccaagaatt ggagccaatc aattcttgcg 11160
gaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgct ccgccatctc 11220
cagcagccgc acgcggcgca tctcgggcag cgttgggtcc tgcagatccg gctgtggaat 11280
gtgtgtcagt taggggtgtg aaagtcccca ggctccccag caggcagaag tatgcaaagc 11340
atgcatctca attagtcagc aaccagggtg ggaaagtccc caggctcccc agcaggcaga 11400
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgcccct aactccgcc 11460
atcccgcccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt 11520
tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga 11580
ggcttttttg gaggcctagg cttttgcaaa a 11611

```

BEST AVAILABLE COPY

<210> 5

<211> 21

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Primer

<400> 5

cgatgtagga gggcgtggat a

21

<210> 6

<211> 21

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Primer

<400> 6

gcttctgcgg gcgatttg t

21

<210> 7

<211> 20

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Primer

<400> 7

tgagaatatc accggaattg

20

<210> 8

<211> 21

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Primer

<400> 8

agctcgacat actgttcttc c

21

<210> 9

<211> 24

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Primer

<400> 9

gtgaatggaa atcccatcgc tgtc

24

<210> 10

<211> 24

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Primer

<400> 10

agtgggtact ctaaaggcca tacc

24

<210> 11

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> .(166)..(1151)

<223>

<400> 11

ggcacgagct tgcacgcaag tcagcgcgcg caagtcaaca cctgccggtc cacagcctca

60

aataataaag agctcaagcg tttgtgcgcc tcgacgtggc cagtctgcac tgccttgaac

120



ccgcgagtct cccgccgcac tgactgccat agcacagcta gacga atg cag cta gca 177  
Met Gln Leu Ala  
1

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225  
Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys  
5 10 15 20

gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273  
Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp  
25 30 35

gag acc cag tac tcg ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321  
Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro  
40 45 50

gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369  
Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile  
55 60 65

aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac 417  
Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His  
70 75 80

gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg 465  
Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp  
90 95 100

ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc 513  
Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser  
105 110 115

ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca 561  
Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu Phe Leu Tyr Thr  
120 125 130

ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg 609  
Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile Ala Met  
135 140 145

BEST AVAILABLE COPY

aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg	657
Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu	
150 155 160	
tac gcc tgg ttt gat tac aac atg ctg cac cgc aag cat tgg gag cac	705
Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys His Trp Glu His	
165 170 175 180	
cac aac cac act ggc gag gtg ggc aag gac cct gac ttc cac agg gga	753
His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp Phe His Arg Gly	
185 190 195	
aac cct ggc att gtg ccc tgg ttt gcc agc ttc atg tcc agc tac atg	801
Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met	
200 205 210	
tcg atg tgg cag ttt gcg cgc ctc gca tgg tgg acg gtg gtc atg cag	849
Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr Val Val Met Gln	
215 220 225	
ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg ttc atg gcg gcc gcg	897
Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala	
230 235 240	
ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt ggc acg tac atg ccc	945
Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Met Pro	
245 250 255 260	
cac aag cct gag cct ggc gcc gcg tca ggc tct tca cca gcc gtc atg	993
His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser Pro Ala Val Met	
265 270 275	
aac tgg tgg aag tcg cgc act agc cag gcg tcc gac ctg gtc agc ttt	1041
Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe	
280 285 290	
ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc	1089
Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro	
295 300 305	
ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga	1137

Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg  
310 315 320

ggt ctg gtt cct gcc tag ctggacacac tgcagtgggc cctgctgcca 1185  
Gly Leu Val Pro Ala  
325  
gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccc 1245  
gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggaggggg tttgttagctg 1305  
tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gactacaccc acaggccaac 1365  
cccttgcag gagatgtctt gcgtcgggag gagtgttggg cagtgtagat gctatgattg 1425  
tatcttaatg ctgaagcctt taggggagcg aacttagtg ctgggcaggc aacgccctgc 1485  
aaggtgcagg cacaagctag gctggacgag gactcgggtg caggcaggtg aagaggtgcg 1545  
ggaggggtggt gccacacca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605  
agagctgcgt gattaactgg gctatggatt gtttgagcag tctcattat tctttgatat 1665  
agatactggt caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725  
ccctgcgctt atgaagctgt aacaataaag tgggtcaaaa aaaaaa 1771

<210> 12

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 12

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala  
1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val  
20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp  
35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp  
50 55 60

Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala  
70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp  
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser  
100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu  
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly  
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val  
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys  
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp  
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met  
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr  
210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe  
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly  
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser  
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp  
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His  
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg  
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala  
325

<210> 13

<211> 1662

<212> DNA

<213> *Haematococcus pluvialis*

<220>

<221> CDS

22> (168)..(1130)

<223>

<400> 13

cggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga 60

gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg 120

ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc 176

Met His Val

1

gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc 224

Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser

5

10

15

agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc 272

Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser

20

25

30

35

gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct 320

Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro

40

45

50

cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc 368

Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly  
55 60 65

acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg 416  
Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro  
70 75 80

aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc 464  
Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala  
85 90 95

cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc 512  
Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe  
100 105 110 115

att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac 560  
Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp  
120 125 130

gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc 608  
Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu  
135 140 145

ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg 656  
Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met  
150 155 160

g cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg 704  
Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly  
165 170 175

aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc 752  
Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe  
180 185 190 195

gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg 800  
Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu  
200 205 210

gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat 848  
Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn

215	220	225	
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu			896
230	235	240	
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala			944
245	250	255	
gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala			992
260	265	270	275
gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp			1040
280	285	290	
gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys			1088
295	300	305	
cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala			1130
310	315	320	
cctgtccct ccgctggtga cccagcgtct gcacaagagt gtcattgctac aggggtgctgc			1190
ggccagtggc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccaactgagca			1250
ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg			1310
ggcgtgctac tgacaatggg cgtgctactg gggctctggca gtggctagga tggagtttga			1370
tgcattcagt agcgggtggc aacgtcatgt ggatgggtgga agtgctgagg ggtttaggca			1430
gccggcattt gagagggcta agttataaat cgcattgctgc tcatgcgcac atatctgcac			1490
acagccaggg aaatcccttc gagagtgatt atgggacact tgtattgggtt tcgtgctatt			1550
gttttattca gcagcagtac ttagtgaggg tgagagcagg gtgggtgagag tggagtgagt			1610



gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct

1662

<210> 14

<211> 320

<212> PRT

<213> Haematococcus pluvialis

00> 14

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala

1

5

10

15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His

20

25

30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala

35

40

45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr

50

55

60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile

65

70

75

80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu

85

90

95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala

BEST AVAILABLE COPY

BASF AG  
BASF NAE 597/03

26/68

4. September 2003

100

105

110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr

115

120

125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu

130

135

140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp

145

150

155

160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly

165

170

175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val

180

185

190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe

195

200

205

Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro

210

215

220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala

225

230

235

240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro

245

250

255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr

260

265

270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp  
275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu  
290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala  
305 310 315 320

<210> 15

<211> 729

<212> DNA

<213> *Agrobacterium aurantiacum*

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 15

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg 48  
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
1 5 10 15

atc gtc tgc gcc gcc atc atc gcc gct tgg ctg gcc ctg cat gtg cat 96  
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
20 25 30

gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala	144
35 40 45	
aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	192
50 55 60	
cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	240
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	288
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	336
100 105 110	
gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	384
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	432
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	480
145 150 155 160	
gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	528
165 170 175	
gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	576
180 185 190	
gac cgc cac aat gcg cgg tcg tcg cgg atc agc gac ccc gtg tcg ctg	624

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
195 200 205

ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac 672  
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720  
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
225 230 235 240

acc gca tga 729  
Thr Ala

<210> 16

<211> 242

<212> PRT.

<213> Agrobacterium aurantiacum

<400> 16

t Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala  
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala  
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn  
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
100 105 110

Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
225 230 235 240

Thr Ala

<210> 17

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99) .. (827)

<223>

<400> 17

ctgcaggccg ggcccgtg ccaatggtcg caaccggcag gactggaaca ggacggcggg 60

ccggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116

Met Ser Gly Arg Lys Pro

1

5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164

Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile

10

15

20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212

Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp  
25 30 35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260  
Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr  
40 45 50

tgg ctg tcg gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308  
Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly  
55 60 65 70

tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356  
Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu  
75 80 85

gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag 404  
Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys  
90 95 100

cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc 452  
His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe  
105 110 115

ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat 500  
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr  
120 125 130

tc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat 548  
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr  
135 140 145 150

gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc 596  
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val  
155 160 165

ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg 644  
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu  
170 175 180

ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg 692



Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg  
185 190 195

tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc 740  
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe  
200 205 210

ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg 788  
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp  
215 220 225 230

cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct 837  
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala  
235 240

cattgtcgtg ggcacagtcc tcgtgatgga gctgaccgcc tattccgtcc accgctggat 897

tatgcacggc cccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc 957

gttgagagaag aacgacctct acggcgctcgt ctccgcgggtg ctggcgacga tcctcttcac 1017

cgtggggcgc tattgggtggc cgggtcgtgtg gtggatcgcc ctgggcatga cggctctatgg 1077

gttgatctat ttcacacctgc acgacgggct tgtgcatcaa cgctggccgt ttcggatatat 1137

tccgcggcgg ggctatttcc gcaggctcta ccaagctcat cgctgcacc acgcggtcga 1197

ggcggggac cactgcgtca gcttcggctt catctatgcc ccaccgtgg acaagctgaa 1257

gcaggatctg aagcggtcgg gtgtcctgcg cccccaggac gagcgctcgt cgtgatctct 1317

gatcccgggc tggccgcatg aaatccgacg tgctgctggc agggggccggc cttgccaacg 1377

gactgatcgc gctggcgatc cgcaaggcgc ggcccgaact tcgcgtgctg ctgctggacc 1437

gtgcgggcgg cgccctcggac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc 1497

actggctgga ccgcctgaag ccgatcagge gtggcgactg gcccgatcag gaggtgcggt 1557

tcccagacca ttcgcgaagg ctccggggccg gatatggctc gatcgacggg cgggggctga 1617

BASF AG  
BASF NAE 597/03

34/68

4. September 2003

tgcggtgcggt gacc

1631

<210> 18

<211> 242

<212> PRT

<213> .Alcaligenes sp.

<400> 18

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu

1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe

20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu

35 40 45

S Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn

65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp

85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr

100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly  
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
145 150 155 160

Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe  
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro  
180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu  
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly  
225 230 235 240

Arg Ala

<210> 19

<211> 729

<212> DNA

<213> *Paracoccus marcusii*

<220>

<221> CDS

<222> (1)..(729)

<223>

<400> 19

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg 48

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat 96

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His

20 25 30

gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg 144

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala

35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

50 55 60

cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat 240

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn

65 70 75 80

gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg 288

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp

85 90 95

cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc 336

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr  
100 105 110

gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc 384  
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
115 120 125

cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc 432  
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140

gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac 480  
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
150 155 160

gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc 528  
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
165 170 175

gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg 576  
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
180 185 190

gac cgc cat aat gcg cgg tcg tcg cgg atc agc gac cct gtg tcg ctg 624  
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
195 200 205

gtg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac 672  
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720  
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
225 230 235 240

acc gca tga 729  
Thr Ala

BASF AG  
BASF NAE 597/03

38/68

4.September 2003

<210> 20

<211> 242

<212> PRT

<213> *Paracoccus marcusii*

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His

20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala

35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

50 55 60

Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn

65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp

85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr

100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala

115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
165 170 175

1 Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro  
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu  
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His  
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp  
225 230 235 240

Thr Ala

<210> 21

<211> 1629

<212> DNA

<213> Synechococcus sp.

<220>

<221> CDS

<222> (1)..(1629)

<223>

<400> 21

g atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta 48  
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu  
1 5 10 15

gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta 96  
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu  
20 25 30

gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg 144  
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met  
35 40 45

ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac 192  
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His  
50 55 60

gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag 240  
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln  
65 70 75 80

tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg 288  
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly  
85 90 95

ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt 336  
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys  
100 105 110



gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa 384  
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln  
115 120 125

ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt 432  
Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe  
130 135 140

aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg 480  
Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp  
145 150 155 160

gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg 528  
Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala  
165 170 175

ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat 576  
Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn  
180 185 190

gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt 624  
Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys  
195 200 205

tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg 672  
Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met  
210 215 220

atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga 720  
Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly  
225 230 235 240

ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa 768  
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln  
245 250 255

ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa 816  
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu  
260 265 270

aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg 864

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg	
275 280 285	
gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg	912
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu	
290 295 300	
caa ttg gtg gaa ccg ggg gcc cta gcc aag gtg aat caa aac cta ggg	960
Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly	
305 310 315 320	
gaa cga ctg gaa cgg cgc act gtg aac aat aac gaa gcc att tta aaa	1008
Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys	
325 330 335	
atc gat tgt gcc ctc tcc ggt tta ccc cac ttc act gcc atg gcc ggg	1056
Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly	
340 345 350	
ccg gag gat cta acg gga act att ttg att gcc gac tcg gta cgc cat	1104
Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His	
355 360 365	
gtc gag gaa gcc cac gcc ctc att gcc ttg ggg caa att ccc gat gct	1152
Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala	
370 375 380	
at ccg tct tta tat ttg gat att ccc act gta ttg gac ccc acc atg	1200
Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met	
385 390 395 400	
gcc ccc cct ggg cag cac acc ctc tgg atc gaa ttt ttt gcc ccc tac	1248
Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr	
405 410 415	
cgc atc gcc ggg ttg gaa ggg aca ggg tta atg ggc aca ggt tgg acc	1296
Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr	
420 425 430	
gat gag tta aag gaa aaa gtg gcg gat cgg gtg att gat aaa tta acg	1344
Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr	

435 440 445  
gac tat gcc cct aac cta aaa tct ctg atc att ggt cgc cga gtg gaa 1392  
Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu  
450 455 460  
agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc 1440  
Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val  
465 470 475 480  
tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488  
Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu  
485 490 495  
cgc gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536  
Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr  
500 505 510  
ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584  
Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg  
515 520 525  
aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629  
Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp  
530 535 540

<210> 22

<211> 542

<212> PRT

<213> Synechococcus sp.

<400> 22

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu

1

5

10.

15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu  
20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met  
35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His  
50 55 60

Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln  
65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly  
85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys  
100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln  
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe  
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp  
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala  
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn  
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys  
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met  
210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly  
225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln  
245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu  
260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg  
275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu  
290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly  
305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys  
325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly

340

345

350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His

355

360

365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala

370

375

380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met

385

390

395

400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr

405

410

415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr

420

425

430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr

435

440

445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu

450

455

460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val

465

470

475

480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu

485

490

495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr

500

505

510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg  
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp  
530 535 540

<210> 23

<211> 776

<212> DNA

<213> Bradyrhizobium sp.

<220>

<221> CDS

<222> (1)..(774)

<223>

<400> 23

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc 48  
Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg  
1 5 10 15

gac gat gcg agg cag cgc cgc gtc ggt ctg acg ctg gcc gcg gtc atc 96

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile  
20 25 30

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg 144  
Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro  
35 40 45

ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg cag	192
Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln	
50 55 60	
acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac	240
Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His	
65 70 75 80	
ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag	288
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln	
85 90 95	
ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc	336
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val	
100 105 110	
gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat	384
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp	
115 120 125	
ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt	432
Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe	
130 135 140	
ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc	480
Leu Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val	
145 150 155 160	
tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg	528
Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu	
165 170 175	
ctg ttc tgg gcg ctg ccc ggg ctg ctg tgc gcg ctg cag ctg ttc acc	576
Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr	
180 185 190	
ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat	624
Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp	
195 200 205	



BASF AG  
BASF NAE 597/03

49/68

4. September 2003

cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg 672  
Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu  
210 215 220

acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat 720  
Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp  
225 230 235 240

gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg 768  
Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg  
245 250 255

gac ta  
Arg Asp

776

<210> 24

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

00> 24

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg  
1 5 10 15

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile  
20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro  
35 40 45

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln

50

55

60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His

65

70

75

80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln

85

90

95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val

100

105

110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp

115

120

125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe

130

135

140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val

145

150

155

160

Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu

165

170

175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr

180

185

190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp

195

200

205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu

210

215

220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp  
225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg  
245 250 255

Arg Asp

<210> 25

<211> 777

<212> DNA

<213> Nostoc sp.

<220>

<221> CDS

<222> (1)..(777)

<223>

<400> 25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta  
Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu  
1 5 10 15

48

ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt  
Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe  
20 25 30

96

att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta	144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu	
35 40 45	
ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc	192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala	
50 55 60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat	240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His	
65 70 75 80	
gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat	288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn	
85 90 95	
ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa	336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys	
100 105 110	
gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat	384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp	
115 120 125	
tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg	432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp	
130 135 140	
tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga	480
Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly	
145 150 155 160	
tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa	528
Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu	
165 170 175	
aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta	576
Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val	
180 185 190	

caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624  
Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly  
195 200 205

ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672  
Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe  
210 215 220

tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720  
Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His  
225 230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768  
Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile  
245 250 255

tct tta taa 777  
Ser Leu

<210> 26

<211> 258

<212> PRT

<213> Nostoc sp.

<400> 26

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu  
1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe  
20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu

35

40

45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala

50

55

60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His

65

70

75

80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn

85

90

95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys

100

105

110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp

115

120

125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp

130

135

140

Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly

145

150

155

160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu

165

170

175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val

180

185

190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly

195

200

205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe  
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His  
225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile  
245 250 255

Leu

<210> 27

<211> 789

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(789)

<223>

<400> 27

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa  
Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
1 5 10 15

48.

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96  
Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
20 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144  
Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192  
Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
50 55 60

atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat 240  
Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
70 75 80

ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca 288  
Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
85 90 95

cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag 336  
Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
100 105 110

aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat 384  
Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp  
115 120 125

ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc 432  
Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe  
130 135 140

atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta 480  
Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu  
145 150 155 160

ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc 528  
Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile  
165 170 175

tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat 576



Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr  
180 185 190

ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat 624  
Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
195 200 205

ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc 672  
Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
210 215 220

gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat 720  
Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
5 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768  
Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
245 250 255

aat tca gta acc aat tgc taa 789  
Asn Ser Val Thr Asn Ser  
260

<210> 28

<211> 262

<212> PRT

<213> Nostoc punctiforme

<400> 28

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
20 25 30

BEST AVAILABLE COPY

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
65 70 75 80

Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp  
115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe  
130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu  
145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile  
165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr  
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile  
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn  
245 250 255

Asn Ser Val Thr Asn Ser  
260

<210> 29

<211> 762

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1) .. (762)

<223>

<400> 29

gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca  
Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro

1	5	10	15	
gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc				96
Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val				
20		25	30	
att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac				144
Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp				
35		40	45	
atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa				192
Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln				
50		55	60	
ca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat				240
Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His				
65		70	75	80
ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca				288
Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr				
85		90	95	
ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa				336
Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys				
100		105	110	
aa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat				384
Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp				
115		120	125	
ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt				432
Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe				
130		135	140	
atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att				480
Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile				
145		150	155	160
tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act				528
Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr				
165		170	175	

tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat 576  
Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr  
180 185 190

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624  
Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln  
195 200 205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672  
Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720  
Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762  
Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
245 250

<210> 30

<211> 253

<212> PRT

<213> Nostoc punctiforme

<400> 30

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro  
1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val  
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp  
— 35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln  
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His  
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr  
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys  
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp  
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe  
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile  
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr  
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr  
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln

195

200

205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile  
210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His  
225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys  
245 250

<210> 31

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3)..(971).

<223>

<400> 31

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47  
Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile  
1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95  
Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu  
20 25 30

tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc 143  
Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala  
35 40 45

cgc gac atc acg cgg ccc aaa gtc tgc ctg cat gct cag cgg tgc tcg 191  
Arg Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser  
50 55 60

tta gtt cgg ctg cga gtg gca gca cca cag aca gag gag gcg ctg gga 239  
Leu Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly  
65 70 75

acc gtg cag gct gcc ggc gcg ggc gat gag cac agc gcc gat gta gca 287  
r Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala  
80 85 90 95

ctc cag cag ctt gac cgg gct atc gca gag cgt cgt gcc cgg cgc aaa 335  
Leu Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys  
100 105 110

egg gag cag ctg tca tac cag gct gcc gcc att gca gca tca att ggc 383  
Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly  
115 120 125

gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac 431  
Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His  
130 135 140

atg acc gtg ggc ggc gca gtg cca tgg ggt gaa gtg gct ggc act ctc 479  
Met Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu  
145 150 155

ctc ttg gtg gtt ggt ggc gcg ctc ggc atg gag atg tat gcc cgc tat 527  
Leu Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr  
160 165 170 175

gca cac aaa gcc atc tgg cat gag tcg cct ctg ggc tgg ctg ctg cac 575  
Ala His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His  
180 185 190



aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg 623  
Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu  
195 200 205

ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc 671  
Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly  
210 215 220

ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg 719  
Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu  
225 230 235

ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg 767  
Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu  
245 250 255

gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg 815  
Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met  
260 265 270

aag cgc ctg aca gtg gcc cac cag cta cac cac agc ggc aag tac ggt 863  
Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly  
275 280 285

ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att 911  
Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile  
290 295 300

cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg 959  
Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp  
305 310 315

tcc aag cgg tag ggtgcggaac caggcacgct ggtttcacac ctcatgcctg 1011  
Ser Lys Arg  
320

tgataagggtg tggctagagc gatgcgtgtg agacgggtat gtcacggctg actggtctga 1071

tggccaatgg catcgcccat gtctggtcat cacgggctgg ttgcctgggt gaaggtgatg 1131

cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc 1191

caggctggcg ttgaatcagt gagggtttgt gattggcggt tgtgaagcaa tgactccgcc 1251

catattctat ttgtgggagc tgagatgatg gcatgcttgg gatgtgcatg gatcatggta 1311

gtgcagcaaa ctatattcac ctagggtgtg tggtaggatc aggtgaggcc ttgcacattg 1371

catgatgtac tcgtcatggt gtgttgggtga gaggatggat gtggatggat gtgtattctc 1431

agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattctt tgagagggga 1491

ggctcgtgcc agaaatggtg agtggatgac tgtgacgctg tacattgcag gcaggtgaga 1551

actgtct cgattgtaaa atacattcag atgcaaaaaa aaaaaaaaaa aaaaaaa 1608

<210> 32

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly

1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser

20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg

35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu

50

55

60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr  
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu  
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg  
100 105 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val  
115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met  
130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu  
145 150 155 160

Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala  
165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys  
180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe  
195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe  
210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly  
225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val  
245 250 255

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys  
260 265 270

G Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly  
275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro  
290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser  
305 310 315 320

Lys Arg